THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR KATHLEEN K. BIERSDORFF

TITLE OF THESIS THE ROLE OF ASSOCIATION RETRIEVAL IN RECOGNITION AND RECALL OF EPISODIC MATERIAL

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED SPRING, 1981

THE UNIVERSITY OF ALBERTA

THE ROLE OF ASSOCIATION RETRIEVAL IN RECOGNITION AND RECALL OF EPISODIC MATERIAL

by

© KATHLEEN K. BIERSDORFF

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Doctor of Philosophy

IN

Psychology

Department of Psychology

EDMONTON, ALBERTA

SPRING, 1981

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled THE ROLE OF ASSOCIATION RETRIEVAL IN RECOGNITION AND RECALL OF EPISODIC MATERIAL submitted by KATHLEEN K. BIERSDORFF in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psychology.

Abstract

Bartling and Thompson (1977) proposed that recognition involves retrieval of backward associations (target to input context) while recall (CRt) involves retrieval of forward associations (input cue to target). Their model assumes that interfering with backward associations reduces recognition while leaving CRt unaffected. Conversely, interfering with forward associations reduces CRt without affecting recognition. Tulving's recognition failure of recallable words is supposedly due largely to the association task creating competing backward associations. Bartling and Thompson experimentally supported this contention but failed to test the corresponding notion that interference with forward associations should reduce recognition failure via reducing CRt. They also assume in their description of recognition processes that retrieval in recognition is like recall of the input cue given the target (CRc). This leads to the prediction that recognition should parallel CRc regardless of which associations experience interference. The first experiment included a partial replication of Bartling and Thompson's third experiment. In addition to the group giving associations to strong associates of the targets (backward associative interference (BAI)), another group gave primary associations to the input cues (forward associative interference (FAI)), and a control group performed a math computation task. The BAI condition had higher recognition failure values compared with the

control. However, the FAI association task tended to result in the target being produced either overtly or covertly. Therefore, a second experiment introduced intervening lists in place of the association task. These lists corresponded to either an A-B;A-C (FAI), an A-B;D-B (BAI), or an A-B;E-F (Control) pattern. FAI recognition failure was reduced with respect to the control but not always via reducing CRt. Furthermore, an interaction between interference condition and recognition vs. CRc resulted from a facilitation in BAI recognition not experienced by BAI CRc. Unlike FAI and Control, BAI subjects saw targets twice, improving recognition but interfering with CRc because of response competition. This suggests that recognition does not necessarily involve retrieval of backward associations, since one can demonstrate interference with backward associations (CRc) while producing facilitation of recognition. Finally, the inviolability of episodic traces suggested by Tulving (1972) was tested. If similar but not identical episodic events are experienced, a retrieval cue (stimulus term of a pair) should have more features in common with the appropriate target than with the inappropriate one. Interference in this situation would suggest that similar albeit not identical episodic traces may interact. Another possible outcome would be that there should be no interference but that wrong answers should favor the similar target. Four lists containing internally associated A-B pairs and one of four types of yoked pairs

(homograph stimuli biased toward the related meaning by the responses, homograph stimuli biased toward the unrelated meaning, homograph stimuli unbiased by unrelated responses, and unrelated internally associated pairs) were presented to separate groups and tested for CRt. No interference was found for A-B pairs but wrong answers favored the similar target in the related homograph list while favoring no response and non-list responses in the unrelated pair list. Under some circumstances, at least, featurally similar traces do not interact.

Acknowledgments

I would like to sincerely thank the members of my committee, whose comments and suggestions have made this a much better thesis. In particular, I would like to thank my advisor, Willie Runquist, whose dedication to common sense and confidence in my abilities made a potentially stressful experience enjoyable as well as intellectually stimulating. Finally, my heartfelt appreciation is extended to Jane Clifton and to my husband, Richard Patching, both of whom untiringly listened to my ideas and never failed to nod where appropriate.

Table of Contents

List of Tables

I. Introduction to the General Problem

The division of memory structure and processes into those involving semantic/knowledge information and those involving episodic information has been recently repopularized (Tulving, 1972). Tulving has made this distinction on the grounds of differences in trace characteristics such as storage processes, longevity of trace, ease of retrievability, and basic organization.

Because the distinction between the episodic trace and semantic memory has become popular recently, several different theoretical approaches to memory have attempted to incorporate episodic memory with several different definitions resulting. Not all of these definitions are compatible. One of the author's intentions is to examine the implications of a few of these alternative definitions both theoretically and experimentally. While it is possible that the utility of the concept itself may come into question, this is not to be considered a primary purpose of this investigation.

Because of the nature of many of the episodic memory tasks, a second issue is relevant. Tulving and his associates (Tulving & Thomson, 1973; Tulving, 1974; Watkins & Tulving, 1975) have often investigated episodic processes via a cued recall - recognition task comparison. Certain suggestions have been made regarding recognition and recall processes that are both basic to the nature of recognition as a one or a two stage process and as well have a direct

bearing on the conceptualization of episodic memory. It is also the author's intention to review some of the arguments surrounding the nature of recognition in comparison with recall and some of the more relevant data. It is hoped that the experimental investigations presented here will shed some light on the nature of recognition and recall as well as on episodic memory.

A. The Recognition Failure Phenomenon and its Interpretations

In 1972, Tulving presented his first formulation of episodic memory. In its simplest form episodic memory is memory for events or episodes. It is autobiographical in terms of the events that are retained, whereas semantic memory information represents concepts and relations that are more general. For instance, each of us learns the basic letter-sound correspondences when learning to read. When we read a sentence we need not recall the circumstances under which we learned the relevant letter-sound correspondences (an episodic trace). The information we require for the task comes automatically from our basic knowledge or, as Tulving would call it, semantic memory.

Another aspect of the episodic-semantic distintion immediately discernable is retrievability of information. It is more than likely that much of the trace for the experience of learning to read has been forgotten over time. However, the ability to use the results of that event is

readily available. Episodic information is much more susceptible to interference and erasure than semantic information. The manner of storage for these traces may also differ. Episodic memories are stored sequentially in terms of the order in which the events occurred. Contiguity here would be in terms of relative time of occurrence. Semantic memories are stored in terms of similar concepts and meaning. Contiguity here would be in terms of cognitive structure. Semantic information is often transformed to fit the existing cognitive structure and may, in fact, be the result not of some external experience but of thought. Episodic memories, on the other hand, are thought to be stored in a veridical fashion.

To this extent episodic and semantic memory are considered to be separate and independent of one another. While semantic memory may use restructuring of existing information to add to its contents, it also relies on episodic experiences to introduce new data. Moreover, episodic events may be defined temporally, but often rely for interpretation on information stored in semantic memory. In this way storage in one system is influenced by the other system. That is, in fact, what Tulving holds responsible for the effect of pre-experimental associations as cues for episodic traces.

With respect to the role of semantic memory information in retrieval of episodic traces, Tulving met with theoretical opposition in the form of the

Generation-Recognition model as interpreted by Bahrick (1970) and others. Essentially this model proposed that recall of information from memory consisted first of generation of possible responses given certain cues, some as simple as instructions to recall a list, some as complex and structured as orthographic, acoustic, and semantically related information. The second stage or process involved in recall consisted of a recognition decision about each of the generated alternatives until the "target" was recognized. The basis of this decision might be strength, familiarity, a list/occurrence tag or some sort of episodic information. Recognition as a task separate from recall was most often held to involve only the latter of these two stages, as the "generation" of a copy of the target and alternatives is experimenter-produced (Kintsch, 1970). In brief, generation-recognition theorists would allow the direct involvement of semantic memory information in retrieval as well as storage. In addition, their position implied more or less perfect recognition of recallable words, allowing for random fluctuations of trace accessability.

Tulving's model made claims that while there might be and usually was semantic memory involvement in the storage of episodic traces, the involvement of the semantic memory system at retrieval is limited to reinstatement of the trace as stored originally, i.e., only those semantic memory features involved in encoding are useful at retrieval. This

position culminated in his statement of the Encoding Specificity Principle (ESP) which states that "What is stored is determined by what is perceived and how it is encoded, and what is stored determines what retrieval cues are effective in providing access to what is stored" (Tulving & Thomson, 1973, p.353). This basic principle has been expanded and delineated many times since, but the basic statement presented here is still the core.

Tulving addressed himself to the issue of recognition as merely a sub-process of remembering in a series of experiments which demonstrated that there were circumstances in which words later recalled were not recognized (Tulving & Thomson, 1973; Tulving, 1974; Watkins & Tulving, 1975) and that these circumstances were relatively robust. This series of experiments presents a phenomenon now known as "recognition failure", which is short for recognition failure of recallable words.

The paradigm in its basic form consists of providing one or more practice lists in which a set of target words is presented, each target accompanied by a weak associate (1%) and then the input cues are presented, each followed by a line on which the subject is to write the word from the list that had been presented with it. After the practice list(s), the critical list is presented. Again a weak associate accompanies each target as an input cue which the subjects are told "might help them at the time of the subsequent test" (Tulving & Thomson, 1973). An example of

this might be "ground COLD". Following presentation of the critical list, the subject would receive a sheet of paper with strong associates of some or all of the target words, e.g., HOT, and be asked to give three free associations to each word. More often than not the target is among them. When the free association task is completed the subject is asked to look over the associations and circle any that were in the preceding list. Then the subject is given the cued recall with the original associates as cues.

The basic paradigm has been modified, among other ways, by leaving off the set-inducing practice lists, by using a forced choice recognition task, by performing the recognition task from among experimenter-provided alternatives, and by reducing the level of similarity among the recognition task alternatives without altering the basic finding that there was a relatively large number of targets recalled that were not recognized (Watkins & Tulving, 1975).

The major point to be made at the time was that one could use context to make normally good cues become relatively ineffective. Simply generating the target or having it generated for one was not sufficient. One of the major assumptions inherent in the generation-recognition model was the transsituational identity of words, i.e., encoding context or situation does not modify the internal representation except by the addition of a list marker, however defined. Therefore, generation of the word in response to a cue generally leads to recognition because the

list marker is there. In Tulving's free association task the target is certainly generated but then fails to be recognized in many cases.

The first response to this result was to claim that the encoding of the target was biased by the unusual (weak) input cue and research showed that, indeed, awareness of the relationships between the target and extra-list cue eliminated the effect (Santa & Lamwers, 1974). In fact, the biasing by context was largely the effect Tulving was trying to produce. When the meaning of the target was changed at encoding, cues that were biased toward a different meaning were ineffective. One could not assume a level of cue effectiveness at retrieval without knowing the conditions of storage.

The end result with respect to this issue was that the generation-recognition theorists modified their model such that a "word sense" rather than a word (copy cue) is what is generated and recognized (Reder, Anderson, & Bjork, 1974).

The phenomenon of recognition failure, by this point, was generating much research both in an effort to find out under what conditions it did and did not hold and what aspect(s) of the task were critical in producing the phenomenon. Beside several modifications such as a forced choice recognition, dropping the set-inducing lists, and providing dissimilar lures in the recognition task, which did not significantly reduce recognition failure, there were some manipulations which drastically reduced the phenomenon.

Leaving out the free association task reduced the proportion recalled but not recognized to within the range of normal test to test fluctuation via increasing the level of recall a little (approximately 10%) and increasing the level of recognition a lot (approximately 25%) (Watkins & Tulving, 1975). Tulving has generally ignored the issue of test to test fluctuation and its effect on recognition failure. It is, first of all, physically impossible to have negative values of $P(\overline{Rn}|Rcl)$ and therefore it should be expected that some level of recognition failure greater than zero might still be considered as not significantly different from zero. Wallace (1978) has looked at factorially combined effects of cued vs. free recall and recognition in a test-retest situation and has found a $P(\overline{\text{1st test}}|\text{2nd test})$ of .066 with repeated free recognition tests and of .058 with repeated free recall tests. Cued recognition followed by cued recognition produced as high a "recognition failure" rate as .181. Rabinowitz, Mandler, and Barsalou (1977) also found a high probability of recognition failure of recognizable words (.17) suggesting that test-retest variability is a problem. Because the conditions that determine the amount of variability are unknown, one is not able to separate out the amount of obtained recognition failure attributable to test-retest variability from that attributable to inability to recognize an item later recalled. In summary, then, it is not reasonable to expect absolute zero recognition failure which suggests that

comparison to an appropriate control is a more reasonable approach than the one normally taken.

Another factor which seems to affect the strength of the recognition failure phenomenon is concreteness. The more concrete the word is, the better cue it is for the remaining pair member. In general, the part of speech is closely correlated with concreteness. Nouns tend to be more concrete than adjectives (although there are a number of exceptions). With respect to the recognition failure phenomenon, the relative concreteness of the target and input cue certainly influences the amount of recognition failure obtained. Using the highly correlated part of speech factor, one would find in general more recognition failure for a noun-adjective combination such as cheese-GREEN than a noun-noun (cheese-CRACKER) or an adjective-adjective (blue-GREEN) combination, and least for an adjective-noun combination such as green-CHEESE (Bartling & Thompson, 1977). In fact, these two effective modifications comprise the data base for an argument by Bartling and Thompson (1977) to explain Tulving's effect and the underlying processes in the cued recall and especially the recognition task.

The differences between Tulving's explanation and Bartling and Thompson's explanation for recognition failure stems from basic differences in the way each views the episodic trace. For Tulving, the episodic trace is an "item-in-context", i.e., the "'item' and 'context' refer to

two different components, both of which are necessary for the definition of an event"(Tulving, 1976, p.68). A literal copy of the target will not be remembered if it does not call to mind the target-as-encoded, which is a complex unit of physical and semantic features unique to that episode. The more recent discussions (Tulving, 1976; Flexser & Tulving, 1978) have looked at the storage situation as a recoding of the target portion of the event in terms of the whole cue-target event. This episodic event should not be thought of simply as an association of two parts, such as the traditional stimulus-response pair, but rather that the meaning of the target is influenced by the input cue in interaction with it(the target). A retrieval cue is effective to the extent that the features that that retrieval cue calls to mind overlap with those of the target-as-encoded (Flexser & Tulving, 1978). It should be noted that this is a non-interference model to the extent that those features that do not overlap contribute nothing negative to the remembering experience. Nor is it the case that any particular featural overlap provides more information in remembering than any other. All features are created equal.

The remembering situation, whether it be some form of recall or recognition, consists of extracting retrieval information from the cues in the retrieval situation and matching this set of extracted features with those in the episodic trace. The item is remembered when the featural

overlap exceeds some criterion presumably set by the person (Flexser & Tulving, 1978). For Tulving, remembering, whether it is recall or recognition, is a two stage process that consists first of an analysis of whatever retrieval information exists in the situation and then a decision process involving the matching of features between the retrieval information and the episodic trace. Recognition failure may be seen as the result of having fewer overlapping features than needed to reach criterion in the recognition situation but more overlapping features than needed to reach criterion in the recall situation.

It may be simpler to see what Tulving is suggesting via an example using the recognition failure paradigm. The pair lamp-TABLE is presented as part of the critical list. The subject has been told that s/he is to remember the word in capital letters and that the word in small letters may help. The subject may then encode TABLE as a complex of features that includes 1) orthographic information about TABLE (and possibly lamp), 2) a flat surface supported by one or more legs, 3) situated in a living room, 4) approximately 1/2m X 1m of surface, and 5) has a lamp on it. The latter three features represent information in the episodic trace that is due to the influence of the input cue of lamp interacting with the coding of TABLE. In addition, one might encode TABLE to conform very precisely to a table with a lamp in one's own living room.

In the free association task the word CHAIR is

presented, and TABLE is given as an association. In the recognition task that TABLE is analyzed according to the circumstances under which it was generated, i.e., in relation to CHAIR. Its features might include 1) orthographic information about TABLE (and possibly CHAIR), 2) a flat surface supported by one to four legs, 3) situated in a dining room or kitchen, 4) approximately 1m X 2m of surface, and 5) has a chair extending from under it. Again, the encoding may be specific to one's own dining room table. A match between this encoding and the episodic encoding, even assuming no information is lost, is made on two features, which may or may not be enough for recognition.

In the cued recall task, lamp is presented. The features it generates, among others, might be 1) the orthographic features of lamp, 2) produces light, 3) situated in a living room, 4) is on a table that is approximately 1/2m X 1m in size. In addition, the lamp may correspond to a lamp on a table in one's own living room. Finally, it must be considered that part of the retrieval information includes the fact that a word related to lamp is being sought. As such, it is likely that the featural overlap in the cued recall situation is both greater than that found in the recognition situation and also greater than the criterion needed for successful recall.

The features chosen for this example may appear arbitrary and the obvious use of imagery in the encodings may be argued with, but whether encoding is imaginal,

propositional, or otherwise represented is unlikely to change the impact of the argument.

Bartling and Thompson (1977), on the other hand, take a slightly more traditional approach to the episodic trace. While they agree that memory tasks all depend on the ability to reinstate the episodic trace, this trace is defined in terms of forward and backward associations between the target and input cue. That is, the episodic trace is the integrated unit created from the input cue and target. The forward and backward associations between cue and target provide the basis of retrieval of the whole.

For Bartling and Thompson, recognition primarily involves using backward associations from target to cue to reinstate the whole, while cued recall primarily involves using forward associations from cue to target to reinstate the episodic trace. Retrieval is conceived here as a directed search. The "better" the cues, the more likely the search will quickly converge on the episodic trace. Unlike Tulving's model, the episodic trace is the product of the search. Recognition failure results when "the subject is able to use the context cue to reinstate the episodic memory trace containing elements from both items in the pair, and the subject is unable to use the target word to reinstate the episodic memory trace involving the pair"(Bartling & Thompson, 1977, p.696). In Bartling and Thompson's (1977) model retrieval or retrieval failure is the mechanism underlying both recall and recognition or failure of an item

to be recalled or recognized. Recognition of the retrieved information appears to be either automatic or irrelevant to their model as stated.

One of the critical elements in the recognition failure paradigm, according to Bartling and Thompson's (1977) findings, is the free association task. Assuming that forward and backward associations are independent (a questionable assumption but a necessary one if the model is to be testable), the free association task differentially interferes with the backward associations by setting up competing backward associations between the target and the word with which it was generated. In the recognition task, then, the person tries to use backward associations from the target copy to reach the input cue element and finds the stronger association between the target and the extralist cue interferes. (One assumes here that the strong extralist cue-target association is as strong in both its backward and forward associations as the weak input cue-target association at least.)

In the cued recall task, however, there is little within the paradigm to compete with the forward associations between the input cue and the target. While their own results (experiment 1) show a slight decrease in cued recall after the free association task, it is not clear as to whether this decrement is due to delay, non-independence of forward and backward associations, natural retrieval asymmetry due to the particular grammatical class

combinations, or some other factor. It is also not clear from the information provided whether this difference is indeed an insignificant one or even what relationship there is between retrieval asymmetry and these items. It seems likely that any decision process that might follow retrieval would not be based on checking the backward association from target to cue for cued recall and vice versa for recognition as the associations disrupted by the free association task would then cause a decrement in the cued recall task similar to that in the recognition task. Rather, the situation more closely corresponds to Horowitz and Prytulak's (1969) concept of a redintegrative memory task in which the part serves to cue the whole and the recognition decision is based on the whole.

A redintegrative memory trace is one in which the input elements form a unit, such as an adjective-noun phrase. It is not necessary to confine the definition to a particular mode of representation such as images or propositions. However, concrete items that are meaningfully related are more likely to have the necessary unitness. Another characteristic of the redintegrative memory trace is retrieval asymmetry. One part of the trace is a more salient cue for the whole than the other. In this way the best cue is most salient in establishing the meaning of the unit. For instance, yellow-CAKE is more similar to chocolate-CAKE than to yellow-BUTTERFLY. The best cue is the one in which uncertainty is most greatly reduced. In

the example above, cake has fewer attributes than yellow has things it may describe. When the most salient cue is missing or another cue substituted for it, redintegration is more difficult than when the less salient portion is missing or replaced.

Note that the Tulving and the Bartling and Thompson lists and paradigm fit remarkably well into this redintegrative framework. The cue-target pairs that produce the most recognition failure are noun-adjective pairs that are meaningfully related. The more salient portion, the noun, easily contacts the episodic trace in the cued recall task. The less salient cue, the adjective, is made even less useful as a redintegrative cue by the creation of competing backward associations or by changing the "meaning" of that item (depending on the point of view) during the free association task. This tends to make the existing retrieval asymmetry even more pronounced.

For Bartling and Thompson, then, recognition failure is a failure in retrieval via competing backward associations, whereas for Tulving it is more a failure to match features. In basic paired-associate terms, for Bartling and Thompson recognition failure is due to inhibition of the B-AB association. For Tulving, it is due to a change in meaning of the B term from study to test.

Tulving's and Bartling and Thompson's views on recognition and recall processes are not the only major views extant. Kintsch (1970) has described recall as

consisting of two stages characterized by retrieval and recognition-decision processes. Recognition as a task consists, in this model, only of a recognition-decision stage. In recognition, retrieval is limited to automatically accessing the corresponding memory representation of the sensorily present item (Kintsch, 1970). The decision of whether the item is indeed a target is based on familiarity, response strength, recency, a list tag, or some other criterion. The actual criterion suggested by two-process theorists seems to be a matter of personal preference rather than data supporting or eliminating any of the criteria.

Recall, on the other hand, does not have the automatic accessing advantage. Retrieval involves trace to trace movements. These trace connections can represent inter-item relationships that come from within the list or may be related in the larger associative network--akin to Tulving's semantic memory. In a similar way, Jones (1978) talks of these two types of relationships as intrinsic and extrinsic knowledge, respectively. Recall is said to be able to make use of either extrinsic or intrinsic information, while recognition is based on intrinsic knowledge relating the target to the context of the presentation.

In brief, then, recall is usually more difficult than recognition because there is an additional place to go wrong, i.e., one may fail to retrieve the target in recall whereas in recognition this is impossible.

The recognition failure phenomenon, then, may be said to pose problems for such a position. However, two process theorists have in general held that the decision criterion may readily change from one situation to another. Kintsch (1978) has suggested that the two stage model is adequate if one allows the possibility that the "subjects employ a considerably less stringent criterion for the editing of implicitly retrieved responses in recall than in a recognition task proper" (p.470). Why they should do so has been a matter of speculation for years. From a variety of sources Kintsch (1978) has found evidence that people may make use of metamemory information such as the ease with which the item was retrieved or the number and likelihood of the response alternatives in setting their decision criteria. Unfortunately, much of this information is unavailable to the researcher prior to the research. While one can experimentally manipulate the decision criterion, one seldom has control of the degree to which it changes, and due to the perversity of experimental subjects one can sometimes not even control the direction of the change. Hence, the predictive value of this position may be rather low.

To summarize, the three positions presented in this paper with respect to the definition of the episodic trace and the processes involved in cued recall and recognition are as follows: 1) Bartling and Thompson (1977) view the episodic trace as made up of a target and input cue forming

an integrated unit. Cued recall consists of using forward associations between the input cue and target portion to retrieve the trace, while recognition consists of using backward associations between the target and input cue portion to retrieve the trace; 2) Tulving and his associates (Tulving, 1976; Flexser & Tulving, 1978) view the episodic trace as the target as influenced by the input cue. Cued recall and recognition both consist of a comparison of the episodic trace to information extracted from the retrieval cues inherent in each situation, with a decision based on the feature overlap between the two; 3) The generation-recognition position as exemplified by Kintsch (1978) and Reder, Anderson, and Bjork (1974) views the episodic trace as a "word sense" suggested by the input cue - target relationship. Cued recall consists of using the cue to search for and generate the response(s) and then a recognition decision based on some form of "familiarity". Recognition involves automatic access to a memory representation (i.e., no retrieval phase) and the recognition decision.

B. Recent Research Bearing on the Theories

There already exists, of course, much data regarding various aspects of these three positions. Resolution of the issues at hand must depend not only on the ability to predict the outcome of the critical tests that this thesis intends to provide, but also must account for the data

already in existence. Therefore, some of the more relevant research will be presented along with the implications for the three positions discussed.

Reder, Anderson, and Bjork (1974) varied the background frequency of the target words in an effort to look at the relationship between the number of different meanings engendered by the word and the specificity of encoding. In their second experiment, high and low frequency targets were presented for learning in the presence of weak associates. Then recognition was tested by means of the Tulving recognition of subject-generated targets and then via a four-alternative forced choice task in which the distractors were of the same background frequency as the targets and also were free associates of the extralist cue-target pair. Finally, cued recall with the input list cue was measured. While the same proportion of targets was generated for high as for low frequency words, 84% of the generated low frequency targets were recognized as compared with only 38% of the high frequency targets. In the forced choice recognition task, the low frequency targets maintained their high level of recognition (85%) and the high frequency targets were more often recognized than before but still not as often (72%) as the low frequency targets. Finally, in the cued recall task, high frequency targets were recalled 61% of the time and low frequency targets 71% of the time.

Reder, et al. (1974) concluded that what was encoded and generated for recognition was not simply a copy of the

target but a particular "word sense". With respect to the present arguments, the data suggest that criterion differences among the various types of recall and recognition tasks is not an adequate explanation for the phenomenon of recognition failure. More specifically, if words that are generated are more likely to have a lower recognition-decision criterion than targets presented for recognition, one might expect the forced choice recognition to produce no more targets than either the cued recall or the generation-recognition task given correction for guessing. It is not entirely clear from their description of the procedure of experiment 2 whether the instructions for circling any targets from among the generated responses were concurrent with instructions to free associate or whether they followed the free association task as is the case with Tulving's previous research on the recognition failure phenomenon (Tulving & Thomson, 1973). If the latter is the case, one ought to consider that the two recognition tasks contain little that would produce criterion differences.

With respect to the Bartling and Thompson model, the forced choice recognition task has fewer elements to interfere with the backward associations than the generation-recognition task. In addition, the alternatives in the low frequency recognition tasks are more likely to tap into associations congruent with the input cue which may in turn elicit the input cue and give rise to the target via

forward associations established in the experiment. In the high frequency recognition tasks the alternatives are naturally, and often designed to be, less likely to elicit the input cue. So, while the forced choice recognition task is less interfering than the generation-recognition task, in the case of the low frequency targets there may even be an active facilitation.

The Reder, et al. (1974) data are also consistent with the current Tulving position which holds that the meaning extracted from the retrieval situation must be similar to that in the episodic trace for successful remembering. In the case of the low frequency targets, there are ostensibly fewer meanings to come in contact with so that the list encoding and the meaning extracted in the recognition task are likely to have many overlapping features. In the case of the high frequency targets, it is impossible to ascertain whether the meaning of the target presented along with strong associates is less likely to be influenced by those associates than when the target is a response to one of the strong associates. It is not an unreasonable assumption but it is not subject to analysis given the form of the data as presented to the reader.

Salzberg (1976) presented evidence of some experimental limitations on the phenomenon of recognition failure as well as suggesting that the specificity of the encoding is dependent on the concreteness of the items being encoded. Salzberg noticed that those word pairs of Tulving and

Thomson (1973) that produced the most consistent recognition failure were noun-adjective pairs in which the target was a property of the cue-noun. The recognition cues tended to be antonyms of the targets. The extralist cue might be less effective because the semantic relationship was incongruous with the encoded relationship. Further, the grammatical class differed. In the first experiment, Salzberg factorially varied the grammatical class of the two cues maintaining the target as an adjective. While the grammatical class per se was not an effective variable, the concreteness of the cue seemed to have an effect. In the second experiment, he factorially varied concreteness and grammatical class of cues controlling for word frequency and cue-target relatedness, both problems in experiment 1, and making use of root nouns and their derived adjectives such as "cloud" and "cloudy" and also root adjectives and their derived nouns such as "dim" and "dimness". For the abstract cues there were no differences in recognition between nouns and adjectives, but recognition was superior to recall in all four cases. Recall of nouns, whether root or derived, was superior to recall of adjectives. For the concrete cues there was a recall-recognition interaction with recognition being superior for root adjectives and their derived nouns and recall being superior for root nouns and their derived adjectives. Further, measures of recognition failure ($P(\overline{Rn}|Rcl)$) showed strong failure for concrete root nouns (.37) and derived adjectives (.28) but weaker recognition

failure, ranging from .11 to .22, for all other cue types. Finally, partialling out the effect of grammatical class, a positive relationship between concreteness and recall -recognition differences, but no relationship with grammatical class when concreteness was partialled out was found.

It appears, then, that Tulving's strong recognition failure effect stems largely from his use of concrete relationships between the input cue and its target, reinforced by the noun-adjective relationship. Salzberg addressed himself in his conclusions to the limits of encoding specificity and characteristics of encoding relevant to the retrieval situation. He suggested that the more concrete the items the more likely that imagery should play an important role in the encoding. If a particular image is encoded, the encoding is therefore more specific than with more abstract, less clearly imageable material. It is certainly the case that an image involving abstract material is possible, but the image comes less readily to mind, is less fixed, and is less likely to be a useful component at retrieval.

Salzberg suggests that "if the encoded representation of the cue-target pair includes a unified image incorporating both cue and target words, then reinstatement of that unified image may be a necessary condition for access to the encoded representation" (p.594). This view is somewhat similar to Bartling and Thompson's (1977) position

which holds that the target and cue are both represented as an integral part of the episodic trace. However, the interaction of the two components in forming an image is still quite consistent with Tulving's position. At retrieval, then, the most salient or central part of the encoded representation would also be the most concrete element. In noun-adjective pairs the noun tends to be the more concrete element (unless it is specifically chosen because of its abstractness). In noun-noun and adjective-adjective pairs the first element of the pair is often taken as the modifier. As well, the concreteness of the two elements is greater separately than when taken as a pair. While "dim" and "black" are both relatively concrete and imageable, the pair "dim-BLACK" is clearly not as imageable as a unit. Because of both of these factors, the target tends to be the more salient cue and recognition is more likely to be superior to recall.

One can also look at the elements of the word pair in terms of the number and strength of associations arouseable within the semantic relationship set up in the cue-target pairing. In the object-property relationship common in Tulving's lists, the object has a limited number of properties. For instance, "cloud" may be described by white, fluffy, dark, rolling, and cumulus, to name a few. "Dark", however, may describe a much broader range of objects. As well, each of these objects is more likely to have more idiosyncratic and less strong a claim on the

modifier. In a traditional generation-recognition approach, the object as a cue generates fewer possible responses than the property as a cue, especially if the semantic relationship is considered. Of course, while this may be a relevant aspect of the situation as the Reder, et al. (1974) results suggest, it is clearly not the only element producing the recognition failure results.

While Salzberg's (1976) evidence places restrictions on the encoding specificity principle (i. e., some encodings are less specific than others), it is not inconsistent with Tulving's view of the episodic trace or recall-recognition differences. The more salient aspect of the trace as a retrieval cue is, of course, more likely to overlap in terms of extracted features with the episodic encoding than the less salient aspect as a retrieval cue. This does not require that the input cue or even the target appear as intact copies in the trace. The interactive nature of the image that Salzberg suggests is often part of the encoding is quite in keeping with Tulving's (1976) view of the episodic trace as "item-in-context". While Salzberg does provide evidence regarding the limitations of Tulving's recognition failure phenomenon, the general model that Tulving has used to explain the phenomenon is not limited to the phenomenon and is not challenged to any great extent by Salzberg's (1976) data.

Salzberg's theoretical views appear to be more closely aligned with Bartling and Thompson's (1977) position.

Retrieval asymmetry in favor of the more concrete member may certainly have an effect in terms of the relative strength of forward and backward associations. And, indeed, Salzberg(1976) appears to hold that the target and cue, even within an interactive image, are elements of the episodic trace. But while Salzberg's theoretical leanings appear to favor the Bartling and Thompson (1977) position, his evidence clearly does not distinguish between Bartling and Thompson's and Tulving's positions in favor of one or the other. About Kintsch's (1970, 1976) position, there is the suggestion that the recall - recognition difference lies in the retrieval phase rather than being due to criterion differences, but no hard evidence to that effect is presented.

Some additional support for the asymmetric nature of Tulving's materials comes from a study by Bowyer and Humphreys (1979). In their second experiment they presented word pairs largely taken from Tulving and Thomson (1973) in either A-B or B-A order with bidirectional paired associate instructions. (A words are Tulving and Thomson's input cues and B words are Tulving and Thomson's targets.) They found that the A words were recognized better than B words and that this was true regardless of whether A was the first or second member of the pair. As well, recognition failure for B words was .27 while recognition failure for A words was only .081. These results suggest that the list characteristics may limit the generalizability of the

recognition failure finding. In addition, as Bowyer and Humphreys did not use an intervening association task, these levels of recognition failure reflect the initial level of retrieval asymmetry inherent in Tulving's materials.

Some tentative support for the notion of directional association involvement in retrieval comes from the fact that not only is recognition of A better than recognition of B, but also recall of B is better than recall of A. This happens despite the fact that the instructions emphasize a need for bidirectional associative encoding. Assume that a pair is asymmetric such that the first half of the pair is more concrete, and thus a better cue, than the second half of the pair. Assume further that recognition and recall both consist of using appropriate directional associations between the information physically present and the pair information not physically present. One would then expect the first half to be better recognized (since recognition may imply implicit retrieval of the remaining portion of the trace) and the second half to be better recalled in a cued recall situation (which is the explicit version of the recognition task just described). The fact that this interaction was obtained regardless of the within-pair order (A-B or B-A) suggests that either the specific pair types used by Tulving and his associates are extremely asymmetric or that the directional association explanation of recognition/recall processes is sound, or both.

In experiment 1 of Bowyer and Humphreys (1979),

confidence ratings were taken in addition to recognition and cued recall measures. While recognition failure for A and B words presented as A-B pairs demonstrated the same results as in experiment 2, when only the highest confidence rating was used, recognition failure dropped to .025 and .073 respectively. Bowyer and Humphreys concluded that since recognition failure did not disappear with a change in criterion, it cannot be attributed to criterion artifact. While this conclusion may not stand up to the test-retest results discussed earlier, it may in a modified form suggest that Kintsch's explanation of recognition failure may be inadequate.

Better evidence with respect to Kintsch's generation-recognition position is presented by Hunt (1975). He looked at the effects that a change in context, similar but not identical to the one in the recognition failure task, has on recognition and on recall. Based on assumptions that confidence ratings for correctly remembered items index the recognition-decision phase and that percent correct reflects more closely the retrieval phase (assumptions inherited from Anderson and Bower (1972)), Hunt suggested that differences produced by context changes in recall and recognition tasks could be traced to the subprocesses involved in each.

In experiment 1, the targets were presented with strongly associated input cues which, it was thought, would encourage specificity of encoding (an interesting albeit not

exactly well-supported notion). The context change was to a weakly associated cue and was made after either the first, second, third, or fourth study trial, or not at all in the control condition. Separate recall and recognition groups were run. The rationale was that if retrieval depends on contacting the correct episodic trace and the recognition decision depends on the amount of contextual evidence associated with the target, then a change in context should reduce the probability of successful retrieval but that the probability of a successful recognition decision should be dependent on the number of study trials.

The results showed that in the recall condition there was a sharp drop in the number of correctly recalled items following a change in context such that the number of prior repetitions did not affect the percentage of correct recall. When there was no context change, percent recalled increased over the four trials. In the recognition conditions a change in context produced a less profound but significant decrease in performance across all four trials but the improvement over trials paralleled that in the no context change group. The confidence ratings for the recall conditions showed the same parallel nature as the recognition performance. Confidence in the correctly recalled responses increased slightly over trials but appeared to be near asymptote by the second trial. The context change appeared to produce slightly less confidence in the correctly recalled responses but apparently this

difference was either not statistically significant or else left untested. Confidence ratings for recognition conditions were taken but were not included in this paper.

On the basis of these results, Hunt suggested that the improvement in recognition performance across trials precluded dependence on information provided by new cues for retrieval of the trace and, therefore, supported an automatic access position. Whatever effect context had in recognition must lie in the decision stage. However, in recall, context clearly affects the ease of retrieval. Because no analysis of differences between the context change and control groups in terms of confidence ratings is presented, it is left for the reader to assume that no significant difference exists and that the change in context produces no decrement in the decision phase of recall. If so, and such tenuous claims as may be based on missing analyses must always be taken with caution, this would rather conclusively deny the tenability of Kintsch's (1970, 1976) claims that decision criterion differences could account for recall - recognition differences. On the basis of the presented data, however, it is possible to claim that decision stage differences are not sufficient to account for all recall - recognition differences.

It is not clear how Hunt (1975) can make the claim that automatic access in recognition is fully demonstrated since albeit there is improvement across trials which is supportive of this notion, there is also a significant

decrement in performance in the context change condition compared to the control group. This suggests that retrieval is somewhat less assured when a different context from the one presented at study exists but that this decrement does not interact with the amount of contextual information available for a recognition decision.

Hunt's (1975) conclusions as he states them are very nicely accounted for by the Bartling and Thompson (1977) position. Hunt claims that his data are evidence for the view that changes in context produce decrements at the retrieval phase of recall but at the decision phase of recognition. Bartling and Thompson could view this as evidence for the differing roles that forward and backward associations take in successful cued recall and in successful recognition. Because there was no free association task to set up competing backward associations, the effect of the context change was merely to force the subject to rely on forward associations between the new cue and old target that are pre-experimentally based. In cued recall this would produce retrieval problems, especially as the new cues are weakly associated with the target. Once retrieved, however, redintegration assures that the item is as confidently reported (or nearly so) as if the original cue was presented.

In the recognition task, on the other hand, retrieval is based on backward associations between the target and cue portion of the trace. The new context should have an effect

via limiting the cues available for successful retrieval of the episodic trace. When the original cue-target pair (no context change) is presented for recognition, the subject may use the input cue's forward associations to retrieve the trace, the target's backward associations to retrieve the trace, or the cue-target combination to automatically contact the trace. Failure to recognize the item is most likely an availability problem and is therefore expected to decrease across trials. When the target is presented with a changed context, at least one less search path is somewhat less assured, but the extent of this loss does not change over trials. This results in parallel recognition performance over trials between the context change and no context change groups.

Hunt's data also fail to provide a significant challenge to Tulving's position. The change of context in the recall task would likely lead to extraction of retrieval information that overlaps very little with the trace. This would not be expected to change much across trials. With respect to the confidence rating data, Tulving's interpretation is less certain. Tulving (1976) dismissed the importance of the confidence data of Anderson and Bower (1972) on which Hunt bases many of his initial proposals. However, Anderson and Bower's confidence judgment was based on the presence of the item on a single trial while recall was based on the total experience of the item on the previous trials. In Hunt's (1975) task the confidence

judgment is based on the same experience as the recall. Therefore, it is possible that Tulving's (1976) objections would not hold here. One might expect, according to Tulving's model, that the number of matching features extracted from the retrieval situation would correspond to the confidence rating. If so, one would expect a drop in confidence to follow a change in context as there is likely to be less overlap in features resulting from using a different cue, but the increase in confidence over trials requires more explanation than the Flexser and Tulving (1978) model provides.

With respect to the recognition data, the effect of the context change is merely to decrease the overlap in features between the retrieval cue information and the episodic trace. The presence of a weakly associated cue may have a somewhat detrimental effect but the presence of the target word appears to provide relatively high overlap. This effect increases with the number of repetitions.

At first, one might expect that the change in context should produce a much lower recognition rate but there are some basic procedural differences which may account for the slight decrement in comparison to that generally found in the recognition failure paradigm. The conditions under which the new cue and the target are first paired differ in important ways. In the Tulving paradigm the target is produced in response to the cue. A pairing or encoding of the elements as a unit is forced. In Hunt's (1975)

experiment the pair are simply presented together. The subject may or may not ignore the presence of the unfamiliar element. Recent studies (e.g., Wenger, Thompson, & Bartling, 1980) have demonstrated that recall of an item facilitates later recognition more than re-presentation of that item. The reason for this (i.e., the effect of generation versus presentation) may be similar to that in the case under discussion. It is certainly the case that being weakly associated, it is unlikely that a generation task could have been used to produce an interactive encoding. Hence, Tulving's model is not really at a loss to explain Hunt's recognition results.

Bartling and Thompson's (1977) research, while not providing the critical evidence to distinguish between their own and Tulving's model, does lay the experimental groundwork for such a critical test. Salzberg (1976) had suggested that the noun-adjective pairing sets up a situation in which there is retrieval asymmetry, that is, one member of the pair is a better retrieval cue for the whole than the other member. While the concreteness of the cue was considered to be the underlying factor, the grammatical class of the target and cue can be used to produce retrieval asymmetry. Bartling and Thompson (1977) suggested that in addition to the fairly high initial level of retrieval asymmetry in favor of the input cue as cue in Tulving's experiments, the free association task also creates or increases retrieval asymmetry in that it provides

competing backward associations from the target. It should be noted that Tulving (1974) has stated that his input cue - target pairs are chosen to produce high levels of recall and low levels of recognition, so that the claim of high initial levels of retrieval asymmetry would not be argued with from that quarter. The effect of the free association task was tested in experiment 1 of Bartling and Thompson (1977).

The cue-target pairs and associates were selected from Tulving and Thomson's (1973) and Watkins and Tulving's (1975) lists. Either the traditional instructions or instructions suggesting that either member of the pair might be viewed as the target were given. Cued recall using either the target or the input cue was tested either prior to or after the free association task. It was hoped that the non-standard instructions would lessen the effect of initial retrieval asymmetry and that the free association task would increase retrieval asymmetry and reduce performance on the post-association recall, especially when the target was used as cue. Finally, because of the reduced retrieval asymmetry with the non-standard instructions, one would expect that the difference between pre- and post-association recall would be greater than for the traditional instructions.

The input cue did, in fact, produce better recall than the target in all cases and the pre-association - post-association difference which was in the predicted direction was larger with the target as cue. One would

expect this since cued recall using the input cue makes use largely of the forward associations between cue and target which are left undisturbed by the free association task while cued recall with the target would mainly require the use of backward associations which are interfered with by the free association task. The small decrement in recall using the input cue after the association task may also be accounted for by time-related loss. The non-standard instructions actually produced better recall than the traditional instructions in all but the post-association recall using the target as cue. As such, the view that equalization of the forward and backward associations would make the loss of backward associations more pronounced an effect when cuing with the target was vindicated. Bartling and Thompson's inability to produce good retrieval symmetry with the non-standard instructions was not taken to be a critical failure and may be traced to such factors as the capitalization of the target words in the input list and the naturally high level of retrieval asymmetry in the Tulving lists.

In experiment 2, Bartling and Thompson verified the basic retrieval asymmetry finding by testing cued recall using the input cue or target as cue for four types of cue-target pair: noun-ADJECTIVE (N-A), noun-NOUN (N-N), adjective-ADJECTIVE (A-A), and adjective-NOUN (A-N). The N-A and A-N pairs produced significant retrieval asymmetry in favor of the noun as cue, while the N-N and A-A pairs

were more symmetric in their retrieval capabilities.

Finally, in experiment 3, Bartling and Thompson reproduced the Tulving recognition failure task with the recognition task being an experimenter-provided free choice task. All four types of pairs produced in the second experiment were used and measures of cued recall, recognition, and recognition failure were taken. Recognition failure as measured by $P(\overline{Rn}|Rcl)$ was highest (.36) for N-A pairs which had been shown to have the strongest retrieval asymmetry in favor of the input cue as retrieval cue. Recognition failure was lowest (.04) for A-N pairs which had been shown to have the strongest retrieval asymmetry in favor of the target as cue. The other two types of pairs fell in between with N-N being .18 and A-A being .08. Recognition followed the opposite course with recognition being lowest for N-A pairs (.61) and highest for A-N pairs (.79). Cued recall was highest for N-A pairs (.66) but dropped to .47 and .44 for N-N pairs and A-N pairs respectively and again to .30 for A-A pairs. There appears to be a strong negative relationship between recognition and recognition failure but no strong relationship between recall and recognition failure. (The corresponding positive relation between recognition and recognition given recall has been also noted by Tulving and Wiseman (1975).)

Bartling and Thompson (1977) have suggested on the basis of their data that retrieval asymmetry in favor of stronger intact forward associations can explain the

recognition failure phenomenon and that the free association task which differentially interferes with the backward associations also produces or increases this retrieval asymmetry and thus the level of recognition failure. Most simply, they hold that recognition failure requires two things: 1) the context cue reinstates the episodic trace including the target, and 2) the target does not reinstate the appropriate episodic trace.

At this point a summary of what the recent literature reviewed has to say about the phenomenon and the three models is in order. The Tulving recognition failure phenomenon appears to be limited to a particular class of inputs, i.e., high frequency, diverse-meaning (Reder, et al., 1974), concrete, related pairs for which there is either initial retrieval asymmetry (Salzberg, 1976; Bowyer & Humphreys, 1979) or retrieval asymmetry induced by a free association task favoring the input cue as retrieval cue (Bartling & Thompson, 1977). These characteristics closely resemble those identified by Horowitz and Prytulak (1969) as a redintegrative memory situation.

Tulving's general model, however, goes beyond his restricted recognition failure phenomenon to define all episodic traces as "items-in-context". The relationship between the target encoding and its context cue may vary according to the semantic memory relationship and the situational aspects of the pairing (e.g., experimenter-presented or subject-produced). As such, the

general model of Tulving (Tulving, 1976; Flexser & Tulving, 1978) has not been challenged seriously by the research presented here.

Bartling and Thompson (1977) have presented research that supports the redintegrative nature of the materials used in the recognition failure experiments. Their suggestion that forward and backward associations are used in retrieval fits well with the existing data but the independence of the forward and backward associations has not received adequate testing as yet.

Bartling and Thompson's (1977) explanation of the recognition failure phenomenon is based, again, on a more general model, as they are able not only to predict recognition failure but also the lack thereof. The use of associations in retrieval is stable across tasks. However, the model does not deal with post-retrieval processes. Either the trace is contacted, in which case it is automatically recognized, or else the trace is not contacted (retrieval failure), in which case there is nothing to recognize.

With respect to Kintsch's generation-recognition model, it would appear that the context change which is generally held to be intimately involved with the recognition failure effect cannot be wholely explained by a corresponding change in the decision criterion (Hunt, 1975; Reder, et al., 1974).

In conclusion, each of the models presented here is an incomplete model of remembering. While certain processes

are analyzed in great detail, others are left unspecified or "automatic" in the sense that there is often no stated rationale as to why or how they occur. For Kintsch's model retrieval is automatic in a recognition task. For Bartling and Thompson the recognition decision is automatic. For Tulving the episodic trace is automatically present for a feature comparison. It seems likely that the automaticity in each case is a matter of lack of proper delineation of the model. Experimentally it is necessary to control the factors which might detrimentally affect the "automatic" processes. It may still be possible to test the validity of the non-automatic portions of the models and either reject the models on the basis of what is readily testable or support the delineated portion of the model and perhaps come closer to delineating the currently "automatic" portions.

While Kintsch's (1970, 1976) explanation of recognition failure in terms of different decision criteria for recall and recognition is problematic both in terms of being relatively untestable as it stands and not predicting some of the already existing data as it is interpreted, a test which discriminates between Bartling and Thompson's (1977) explanation and Tulving's (Tulving, 1976; Flexser & Tulving, 1978) explanation should be attempted.

However, Bartling and Thompson (1977) have made some fairly specific assumptions regarding independence of associations and specific sub-processes involved in recognition and recall. Since the viability of a model

sometimes depends on the assumptions it makes either explicitly or implicitly, the direct comparison of the two models was delayed in favor of testing the validity of some of the assumptions of Bartling and Thompson's model.

II. A Test of the Validity of Bartling and Thompson's Model

Bartling and Thompson (1977) assume that forward and backward associations are independent and, therefore, only the backward associations are disrupted by the free association task. In their major demonstration of the recognition failure phenomenon controlling the initial level of retrieval asymmetry (experiment 3), they fail to provide a test of this independence assumption. In order to ascertain whether there is interference with the backward associations while the forward associations remain intact, one must compare retrieval of both A and B terms using the same task. If the associations are independent, one should find that cued recall of the A terms is depressed in comparison to cued recall of the B terms, or in comparison to a group that does not perform the free association task. In experiment 1 which did ask for recall of both A and B terms, it was difficult to determine what was responsible for the decrease in recall of the targets after the association task. The situation was complicated by the use of lists mixed with respect to retrieval asymmetry and the physical features of the to-be-learned elements.

As well, the assumption regarding the independent disruption of directional associations implies something more general about retrieval failure. Not only must one be able to interfere with backward associations without affecting corresponding forward associations, but one must also be able to interfere with forward associations without

affecting corresponding backward associations. Bartling and Thompson (1977) attempted to test only the first half of this more general assumption. It may be possible to demonstrate not only recognition failure when the backward associations are disrupted, but also an increased level of recall failure when the forward associations are disrupted. If so, this would further support Bartling and Thompson's model of retrieval processes but is not critical to the model's success. However, Bowyer and Humphreys (1979) have demonstrated that in a Tulving-type paradigm, the level of recognition interacts with the level of recall. Recognition primes recall so that words that are recognized have a greater probability of being later recalled than those not recognized. The priming effect may act to compensate for the disruption of the forward associations in this case.

Bartling and Thompson's third experiment can be modified to provide a test of both questions. Instead of having cued recall for only the target given the input cue, one would have cued recall for both the targets and the input cues. In order to measure the degree to which the specific direction of associations is disrupted (and also control the study effect of the previous recognition task) a control group would follow an identical procedure except that the free association task would be replaced with a continuation of the computational task used to control for recency. An additional difference in procedure required to test their model of retrieval involves disrupting the

forward associations rather than the backward associations by asking subjects to give what they believe to be the most common response to a list of words that includes the input cues from the critical list. Asking for the normative response should increase the probability that the targets do not appear among the responses, which would reinforce rather than disrupt the forward associations. Finally, since the critical assumptions do not require all four grammatical combinations in order to be tested, only those sets of pairs in Bartling and Thompson's list having initial retrieval asymmetry (A-N and N-A) need be used. In Bartling and Thompson's second experiment they demonstrated that A-N and N-A pairs had initial retrieval asymmetry in favor of the noun (N) as cue. In the recognition failure paradigm (experiment 3) the result of disrupting the backward associations was to create a situation favoring the input cue as retrieval cue. The net effect was to increase the retrieval asymmetry in the pairs for which initial retrieval asymmetry favored the input cue as retrieval cue, i.e., the N-A pairs. As well, for the A-N pairs which had initial retrieval asymmetry favoring the target as retrieval cue, the free association-produced asymmetry favoring the input cue tended to balance it out. Recognition failure for the A-N pairs was very close to zero. Thus, one would predict that if the forward associations are disrupted and the target is now favored as retrieval cue, then retrieval asymmetry should be greatest for the A-N pairs and should

balance out the initial retrieval asymmetry in the N-A pairs.

It is to be expected that if Bartling and Thompson's model holds, one would find an interaction between group (forward associative interference and control) and type of cued recall (recall of the input cue and recall of the target). As well, one would expect post-association differences in retrieval in favor of the target as cue to be greatest in the situation in which the forward associations were weakest to begin with, i.e., the A-N pairs. While the efficacy of the input cue in the N-A pairs should be reduced after the free association task relative to its pre-association task strength as measured in the control group, its efficacy should be somewhat equal to that of the target as cue, that is, retrieval asymmetry should be reduced. In addition, recognition should be better for the A-N pairs than for the N-A pairs and recognition performance should be fairly high in both cases. If forward and backward associations are independent and only forward associations are affected by the free association task, then recognition performance should be the same in the experimental group who performed the free association task as in the control group who did not. Performance in the backward associative interference group, which is the basic Tulving paradigm with the addition of a cued recall of the input cue, should replicate the major findings in Bartling and Thompson's third experiment.

A. Method

Design

The first experiment is a 3 X 2 X 3 design with repeated measures. The three variables are group, pair type (A-N and N-A), and measure (recognition (Rn), cued recall of the input cue (CRc), and cued recall of the target (CRt)). The Forward Associative Interference (FAI) group was asked to give normative associations to the input cues. The Backward Associative Interference (BAI) group was asked to give normative associations to words which have the targets as primary associations (a replication of Bartling and Thompson's (1977) experiment 3). The control group had no association task. Both N-A and A-N pairs were tested for the experimental and control subjects. Measures of recognition, cued recall of half the input cues given the target, and cued recall of the other half of the targets given the input cue were taken from each subject in all three groups. Counterbalancing of the order of the cued recall tasks was not considered necessary because the material being cued in the two tasks did not overlap. However, the items being cued were counterbalanced such that for each pair the input cue was cued half of the time and the target was cued the other half of the time.

Subjects

Subjects for this experiment were 60 undergraduates at the University of Alberta who participated in order to

fulfill an introductory psychology course requirement. These subjects were randomly assigned to one of the three groups on the basis of order of appearance at the lab such that no more than five subjects in a row were assigned to the same condition. Subjects were run individually or in groups of two or three in order to maximize the probability that the instructions were understood and followed. One subject in the FAI condition who gave the target word as a normative association to the input cue for more than 4 of the 24 items was replaced as the experimental hypotheses were based on the situation in which a non-target is generated in the association task. Those cases in the FAI condition in which a target was generated in response to the appropriate input cue were not included in the analyses for that reason, and the analyses were all based on proportions of included items.

Materials

A practice list preceded the critical list. The practice list consisted of the 12 N-N and 12 A-A pairs used in Bartling and Thompson's (1977) third experiment plus four primacy buffer pairs. The critical list consisted of the 12 N-A and 12 A-N pairs from that same experiment except for 2 A-N pairs whose input cues were changed because they were strong forward associates of other targets. Again, four primacy buffer pairs were added which were different from those in the practice list. In general, the target words were low frequency associates (approximately 1%) of the

input cues (Palermo & Jenkins, 1964; Jenkins, 1952) and were moderate to high frequency words according to the Thorndike and Lorge (1944) norms. The critical pairs are presented in Table 1. The critical list pairs were typed on a slide with the input cue having only its initial letter capitalized typed above the target which was completely capitalized.

A test booklet consisting of 10 pages for the FAI and BAI groups and 9 pages for the control group with a blank sheet between each was provided for each subject. For both the experimental groups and the control group the first and fifth pages contained instructions for the practice and critical list study trials. Pages 2 and 6 contained math problems for each subject. The experimental groups had brief instructions on page 7 for the normative association task and the 28 critical list input cues (FAI) or primary stimuli for targets (BAI) plus some unrelated filler stimuli with a line beside each for the normative response.

On page 8 of the two experimental group booklets and page 7 of the control group booklets were free choice recognition instructions followed by 24 rows containing one target and three distractor items each. Pages 3 and 9 of each experimental booklet (and pages 3 and 8 of each control booklet) contained brief cued recall instructions for the input cues given the target and below them six targets from each of the two pair types (A-N and N-A). Pages 4 and 10 of each experimental booklet (and pages 4 and 9 of each control booklet) contained brief cued recall instructions for the

Table 1

Input Cue-Target Pairs in the Critical List for Experiment 1

N-A		A-N	
Input Cue	Target	Input Cue	Target
Cabbage	ROUND	Red	HAIR
Stem	SHORT	Yellow	CAB
Home	SWEET	Dark	ALLEY
Stomach	LARGE	Swift	HORSE
Train	BLACK	Loud	HORN
Country	OPEN	High	LADDER
Butter	SMOOTH	Heavy	FOOT
Glass	HARD	Slow	TURTLE
Cheese	GREEN	Hungry	BEAST
Ground	COLD	Beautiful	PICTURE
Cave	WET	Foggy*	CLOUD
Barn	DIRTY	Crisp*	PICKLE

*Input cues not included in Bartling and Thompson (1977)

target given the input cues below which were the untested pair input cues each followed by a line for the subject's response.

Procedure

Subjects were provided with a written set of instructions and asked to read along with the experimenter's verbal instructions. The instructions indicated that they would see a list of word pairs both terms of which they were to try to remember as either might be tested later.

The sequence of the two experimental condition procedures was as follows with instructions at appropriate points: a) presentation of the 28 pair practice list at a 3-sec. rate via a Kodak Carousel projector, b) a 1 minute math computation task, c) a 2 minute cued recall of input cues given targets, d) a 2 minute cued recall of targets given input cues, e) presentation of the 28 pair critical list at a 3-sec. rate via slides, f) a 1 minute math computation task, g) a 3 minute normative association task using the input cues as stimuli (FAI), or words having the target as primary response as stimuli (BAI), h) a 3 minute recognition task for the targets, i) a 2 minute cued recall of critical list input cues given targets, and j) a 2 minute cued recall of critical list targets given input cues.

The sequence of procedures for the control condition was the same as for the experimental conditions except that f) and g) were replaced by a 4 minute math computation task.

B. Results and Discussion

Noun-Adjective Differences

Unless otherwise stated, statistical results are considered to be significant at $p<.01$ due to the large number of analyses performed here. The supremacy of the noun as cue is without a doubt the strongest effect in the first experiment, as can be seen in Table 2. This, of course, was not unexpected. The purpose of this study was not to cancel the effect but merely to manipulate this effect in ways which were predictable from the model. There are very clear differences between the two types of word pairs demonstrating that when the noun serves as cue, as in the case of cued recall of the target for N-A pairs and recognition and cued recall of the input cue for A-N pairs, the mean percent correct is higher than for the corresponding conditions in which the adjective is cue.

The main analysis of variance summary shown in Table 3 verifies this difference. This analysis involves only the directly obtained measures of Rn, CRc, and CRt. The derived measures (recognition failure, recall failure, and retrieval asymmetry ratio) are analyzed separately in connection with the assumptions and predictions of the Bartling and Thompson model. Otherwise, all of the levels of the three factors described in the design section are fully represented. While the pair type itself was not significant, the measure by pair type interaction was extremely strong ($F(2,114)=187.62$). By referring back to the means in Table

Table 2

Mean Percent of Items Remembered, Recognition Failure, Recall Failure, and Retrieval Asymmetry Ratios

		Rn	CRc	CRt	$\%(\overline{Rn}\mid CRt)$	$\%(\overline{CRt}\mid Rn)$	RAR
A-N	Control	65.4	52.5	36.2	7.0	47.7	.43
	BAI	56.1	50.0	49.3	17.0	31.3	.54
	FAI	57.6	64.8	43.6	20.0	38.7	.39
	Total	59.7	55.8	43.0	14.7	39.2	.45
N-A	Control	51.4	31.6	65.8	34.0	29.1	.61
	BAI	34.6	28.3	73.3	64.4	15.8	.81
	FAI	44.8	41.4	65.4	37.0	21.5	.65
	Total	43.6	33.8	68.2	45.1	22.1	.69

Rn=%Correct Recognition

CRc=%Correct Cued Recall of the Input Cue Given the Target

CRt=%Correct Cued Recall of the Target Given the Input Cue

$\%(\overline{Rn}\mid CRt)$=%Recognition Failure of Recallable Targets

$\%(\overline{CRt}\mid Rn)$=%Recall Failure of Recognizable Targets

RAR=Retrieval Asymmetry Ratio=#CRt/(#CRt + #CRc)

Table 3

Main Analysis of Variance Summary Table

for Experiment 1 Recognition and Cued Recall Measures

Source	SS	df	MS	F
Group(Cntrl vs BAI vs FAI)	1126.24	2	563.12	<1
S(Group)	77967.76	57	1367.86	
Pair type(A-N vs N-A)	1955.34	1	1955.34	2.73
Pair type X Group	195.07	2	97.54	<1
Pair type X S(Group)	40891.92	57	717.40	
Measure(Rn vs CRc vs CRt)	7809.74	2	3904.87	6.54*
Measure X Group	8329.34	4	2082.35	3.49*
Measure X S(Group)	68110.71	114	597.46	
Measure X Pair type	37823.73	2	18911.87	187.62*
Measure X Pair X Group	386.08	4	96.52	<1
Measure X Pair X S(Group)	11491.36	114	100.8	
Total	256087.33	359		

*$p < .01$

2 one can identify the effect as being due to the noun-as-cue superiority. As well, the measure by group interaction was significant ($\underline{F}(4,114)=3.49$) which indicates that the superiority of one group over another depends on the particular measure used. For instance, for the A-N pairs in Table 2 one can see that the control group is superior to the BAI group on recognition but the reverse is true for cued recall of the target. Finally, there were differences among the measures themselves ($\underline{F}(2,114)=6.54$) which reflect the general inferiority of the cued recall of the input cue and superiority of the cued recall of the target which are modified by the higher order interactions. The simple effects inherent in these data will be discussed later in relation to the specifics of the Bartling and Thompson (1977) model.

The mean percent of targets generated in the BAI word association task was 59% with an average of 77% of targets generated for N-A pairs and 40% for A-N pairs. The difference here is probably due to the ability to use the antonym relationships inherent in many of the adjective targets but not the noun targets. As well, one can predict with less effort and more confidence the primary response for adjectives than one can for nouns in many cases. In the FAI word association task in which the intention of the author was to produce non-target competing responses, intrusions of the targets were few (5% for N-A and 2% for A-N pairs) but again appeared to be partly related to

differences in ease of identifying primaries for nouns and adjectives. This will be discussed at greater length later.

Assumptions and Predictions of the Bartling and Thompson Model

Bartling and Thompson's (1977) model suggests that Tulving's paradigm increases the retrieval asymmetry in favor of the noun as cue both by the choice of within-pair relationships (initial asymmetry) and via the effect of the association task used on the backward associations. Their model incorporates several assumptions which have been tested here: 1) that forward and backward associations are independent, 2) that forward and backward associations are involved in specific ways in retrieval, and 3) that recognition is in some way equivalent to an implicit cued recall of the input cue given the target. They suggest that their (and Tulving's) experimental manipulation is effective via disrupting the backward associations between target and input cue.

Because many of their predictions involve comparisons of pairs of groups on a single measure or comparisons of pairs of measures for a single group, it was felt that individual analyses involving only the groups or measures of interest would be more appropriate to test these specific predictions. The results of these individual analyses may suggest that there are problems with some of these assumptions and interpretations.

Independence of Directional Associations

The independence assumption suggests that there should be an interaction between cue type and FAI vs. Control and between cue type and BAI vs. Control as each interference group should experience a drop in performance for the cued recall type whose necessary associations are affected and no difference for the other cue type whose necessary associations are unaffected. Two separate analyses of variance were performed in connection with this prediction. The first compared the Control and FAI groups on the measures of CRc and CRt (cue type) summed across both pair types. The second analysis compared the Control and BAI groups in the same manner. Neither analysis showed the significant interaction expected ($\underline{F}(1,38)=1.39$, MSerror=1198.16 for FAI vs. Control and $\underline{F}(1,38)=3.22$, MSerror=846.44 for BAI vs. Control). (Analysis of variance summary tables for these data can be found in Appendix A.) As well, the cue type that was expected to be comparable to the Control (cued recall of the target (CRt) in the case of BAI and cued recall of the cue (CRc) in the case of FAI) in both cases showed some improvement compared to the Control group. While the increase is not significant, the pattern of the results indicates that the independence assumption is unsupported here.

Involvement of Directional Associations in Retrieval

The evidence for the involvement of directional associations in retrieval is somewhat mixed. Separate

analyses of variance were performed involving only the recognition measure summing over pair type. One analysis compared the recognition performance of the BAI with the Control group. The other compared the recognition performance of the FAI with the Control group. The evidence from the recognition data, also in Appendix A, suggests that the association task used in the Tulving paradigm (BAI condition) does impair recognition to some extent ($\underline{F}(1,38)=6.46$, MSerror=1050.42, $\underline{p}<.02$) when compared with the Control condition, which would be expected if intact backward associations are necessary for recognition. As well, there was no significant decrement in recognition for the FAI group compared to the Control ($\underline{F}(1,38)=1.78$, MSerror=1161.85). This result also fits the conception of backward associations being necessary for recognition.

According to the model, the recognition failure data as measured by $\%(\overline{Rn}|CRt)$ obtained for each subject, show an increase for the BAI condition when compared to the Control. As expected, there was an increase in recognition failure which was statistically significant for N-A pairs but not for A-N pairs using a Newman-Keuls analysis. An analysis of variance performed on the recognition failure data of each group for each pair type and reported in Appendix B verifies the difference in recognition failure in favor of the N-A pairs ($\underline{F}(1,57)=16.99$, MSerror=1550.6), which replicates the Bartling and Thompson (1977) results. There were also significant differences among the groups ($\underline{F}(2,57)=97.68$,

MSerror=43.26) with the BAI group showing the most recognition failure. The interaction between group and pair type was not statistically significant.

The recall failure data, as measured by $\%(\overline{CRt}|Rn)$ obtained for each subject, failed to produce the hoped for increase for FAI in comparison to the Control. This result is not critical to an analysis of the model given Bowyer and Humphreys' (1979) demonstration of priming effects in a similar situation. Priming effects, as well, may be responsible for the decrease in recall failure in the BAI condition compared to the Control. An analysis of the recall failure data of each group for each pair type, also reported in Appendix B, confirmed the difference between A-N and N-A pairs ($\underline{F}(1,57)=7.52$, MSerror=1163.92) this time, naturally enough, in favor of the A-N pairs.

The data from the FAI condition appear to give evidence against the involvement of directional associations in retrieval. While a decrease in recognition failure was predicted by the model, an increase was obtained. As well, no increase in recall failure was obtained, although this is not critical to the model's success. The most likely reason is that the FAI manipulation did not have its intended effect. However, this will be discussed in more detail later.

<u>Recognition as an Implicit Cued Recall of the Input Cue</u>

The suggestion that recognition is equivalent to an implicit cued recall of the input cue received mixed support

in these data, as presented in Appendix A. An analysis of variance was performed comparing Rn and CRc measures for the three groups summed across pair type. There was a marginally significant effect for measure ($\underline{F}(1,27)=4.96$, $MSerror=1138.21$, $\underline{p}=.034$) but no interaction between measure and group. If recognition could be reasonably thought of as an implicit backward cued recall, one would expect the same general pattern for each group in terms of the relative levels of Rn and CRc. An interaction would imply that there are some circumstances in which recognition should not be thought of as an implicit cued recall of the cue given the target. The lack of a significant interaction here provides some support for Bartling and Thompson's assumption (to the extent that one can claim support for a null hypothesis).

The marginally significant effect for measure suggests here that there may be some differences between the two measures which more than likely do not represent a qualitative difference in terms of the processes involved. The fact that the cued recall is reduced with respect to the recognition suggests that the differences cannot be attributed to test-retest effects. Criterion differences suggest themselves as an alternative. It is possible that the retrieval processes are the same but that the decision criterion used for the two tasks differ. The Tulving position, of course, includes this as a possible locus of task differences but allows, in addition, for the two situations to provide differing retrieval information for

feature extraction.

<u>The Role of Retrieval Asymmetry in Recognition Failure</u>

The main consideration of this study concerned the retrieval asymmetry, association task effects on it, and the proposed role of associative interference in it. A summary of analyses of variance related to these issues is contained in Appendix C. The initial asymmetry favoring the noun as cue is well established for all measures in the control group, as demonstrated in the significant Pair type X Cued recall type interaction ($\underline{F}(1,76)=19.82$, MSerror=568.06). As can be seen in the means for the Control condition in Table 2, presentation of the noun for recognition or as the retrieval cue leads to better remembering than the corresponding case in which the adjective is presented for recognition or serves as the retrieval cue. The BAI condition replicates the standard Tulving results and Bartling and Thompson's (1977) experiment 1 results.

The retrieval asymmetry ratio (RAR), as measured by the number of correct CRt divided by the sum of correct CRt and CRc, reflects the relative cuing efficacy of the target and input cue to reinstate the other pair member. If neither pair member is a better retrieval cue, then there is retrieval symmetry reflected in a RAR of .50. If the target is a better retrieval cue than the input cue, then a value of greater than .50 will be obtained. If the input cue is a better retrieval cue than the target, then a RAR of less than .50 will be obtained. Because RAR is a proportion, its

theoretical limits are 0 and 1.

The Control condition, which represents the initial asymmetry level for the word pairs, demonstrates the noun superiority in a RAR value of greater than .50 for the N-A pairs (specifically .61 as presented in the last column of Table 2) and less than .50 for the A-N pairs (specifically .43). The strengthening of this retrieval asymmetry by the association task in the case of the N-A pairs is demonstrated by comparing the corresponding RAR values of the BAI condition (.81) with the Control condition. In the case of the A-N pairs, the presence of the association task would be expected to produce higher RAR values (closer to the .50 of retrieval symmetry) via disrupting the backward associations of the naturally stronger noun cue functioning in the CRc measure. This increase, in fact, was obtained. The retrieval asymmetry ratios for Control and BAI groups reiterate the success of the replication of Bartling and Thompson's (1977) research. The analysis of the RARs for BAI when compared to the control and the appropriate shift in the RAR for BAI when compared to the control show a significant shift in RAR in the expected direction ($\underline{F}(1,38)=8.75$, MSerror=519.40).

The most serious problem is that the FAI condition failed to produce the effect predicted in the Bartling and Thompson model. The presence of competing forward associations should reduce the RAR relative to the Control via reducing the level of CRt while not affecting

recognition or CRc. The analysis of RARs for the Control versus FAI group showed a non-significant decrease for A-N pairs but an equal and opposite effect for the N-A pairs. Pair type itself was the only significant factor ($F(1,38)=11.69$, MSerror=828.26). Two conclusions are possible. Either the procedural mirroring of the BAI and FAI conditions is overshadowed by extraneous factors entering into the FAI condition, i.e., the manipulation did not have the desired effect, or the directional associative interference explanation of recognition failure is not supported.

Problems with the FAI Manipulation

There is some circumstantial evidence suggesting there were unexpected differences between the BAI and FAI procedures. While conducting the experiment it could be seen that the BAI group had no difficulty completing the association task. Everyone appeared to be able to finish the task well within the 3-minute limit and no one in that group complained of the difficulty of the task or any particular items.

The FAI condition association task proved a different matter. Many of the subjects were writing responses until the end of the time limit. (A few filler items at the end of the word association list alleviated the problem of incomplete protocols.) On more than one occasion subjects complained of the difficulty of the task saying that the list response kept coming to mind and that they knew it was

not the most popular response and therefore incorrect. The fact that one or two list intrusions were not uncommon attests to the difficulty encountered. It seems as though there was interference produced, but during the association phase rather than the retrieval phases of the experiment. While the interference is indeed seemingly due to competing forward associations, the placement of the interference within the experiment restricts the utility of direct comparisons. An additional factor is that the majority of the intrusions were in response to nouns (12 as compared with 5 for adjectives). As many of the adjectives had natural opposites while the nouns seldom did, the association task components may have varied in the extreme between the two pair types. The results themselves appear to reflect this discrepancy. The retrieval asymmetry ratios suggest that the interfering effect on the N-A pairs was not as great as expected. In addition, the simple effects analysis of cued recall type at N-A and at A-N suggests that the failure to achieve the desired effect rests mainly in the A-N pairs.

It is likely that the imbalance in intrusions reflects the tendency to implicitly (perhaps consciously) respond with the list response and feeling it to be a non-primary association to attempt to come up with a different-from-list response. Unlike the adjective stimuli which tended to produce adjective responses that were often antonyms rather than nouns (with colors being a possible exception), the

noun stimuli tended to produce adjective responses that were often related to both the noun stimulus and the list response. For instance, GLASS would be more likely to lead to a response of SMOOTH than of WATER or WINDOW. The type of relationship between stimulus and response would be maintained from study trial to association task. As well, a claim for mediation operating here might be made. In addition, there was some re-pairing of list responses in the association task as suggested by the previous example.

In summary, then, there is some evidence to indicate that the association task as it is used in the FAI group may serve as an additional study trial. Tulving and Watkins (1974) cited several studies in which a single trial on an A-B list is followed, without an immediate A-B test, by a single A-C trial. Under those circumstances, subjects do better on A-B and worse on A-C than when there is immediate testing of A-B. Presumably the lack of an immediate A-B test serves as a signal to maintain the A-B association (which also impedes A-C learning). The evidence in the present situation suggests that the FAI association task leads to elicitation and maintenance of the A-B association while attempting to produce a different-from-B association for the task at hand.

The elicitation of the list target during association is also more prominent for the N-A pairs than for the A-N pairs. The occasional re-pairing of items during the association task, also differentially favoring the N-A

pairs, may result in backward associative interference. Any or all of these could be responsible for the discrepancies between expected and observed recognition failure, RAR, and percent recognition and cued recall.

The evidence suggests that certain assumptions of the Bartling and Thompson model are weak. There is little support here for the notion that one can interfere with backward associations while leaving forward associations unaffected. It is also possible that there are some situations in which recognition may not behave like an implicit cued recall of the input cue. However, the obvious problems with the FAI condition make further conclusions about Bartling and Thompson's assumptions difficult. Therefore, it is necessary to make an attempt to eliminate the problem of extraneous factors entering the interference-producing phase of the experiment.

III. An Alternative Approach: Experiment 2

The major problem is the equalization of the interfering tasks in terms of basic difficulty without a change in the amount of or type of interference the tasks produce at retrieval time. An alternative to the word association tasks is an intervening paired associate list. This procedure has the advantage that it allows the equivalent of 100% generation in the standard recognition failure paradigm (BAI) and 0% intrusions in the FAI condition. Thus, if the original list is represented as A-B pairs such as Cheese-GREEN, the FAI intervening list would include A-C pairs such as Cheese-CRACKERS, the BAI intervening list would contain D-B pairs such as Grass-GREEN, and the Control intervening list would be entirely composed of E-F pairs unrelated to the original list pairs.

There are not many studies in the literature using this paradigm so that the potential problems with this sort of task are not well known. Unlike the current study, much of the paired associate research has dealt with arbitrary pairs and often involves multiple learning trials and a single type of retrieval task. However, several unpublished studies reported by Tulving and Watkins (1974) have looked at single trial transfer using non-arbitrary pairs. The purpose of these studies was to look at the effect of testing or not testing the first list on learning the second list associations and retrieving both response sets. Even so, certain characteristics of the paradigm become evident

which are relevant to the issue at hand, namely the control of the type and amount of interference produced.

The first of the studies discussed by Tulving and Watkins (1974) (Tulving, Patterson, & Malis) used weakly associated pairs in an A-B;A-C paradigm with an A-B;C-D control. No test of the A-B list was made prior to presentation of the second list. Recall of both response sets was tested afterward using free recall and modified-modified free recall (MMFR) procedures. They found that there was better recall on the first (A-B) list than on the second list for the A-C condition but not necessarily for the C-D condition. They named this the "priority effect" because the first-learned responses to the stimulus have priority. Malis' unpublished dissertation (also reported in Tulving and Watkins (1974)) looked at the effect on priority of testing neither or both lists immediately after presentation of each. He found that the priority effect on the later MMFR disappeared when the lists were tested immediately after presentation.

Tulving and Watkins (1974) themselves looked at the effect of testing one or the other list immediately after presentation as well as replicating Malis' A-B;A-C conditions on the later MMFR results. In their experiments Tulving and Watkins used unrelated words, but the actual probabilities of correct recall did not appear to vary significantly from those obtained by Tulving, Patterson, and Malis who used weakly associated pairs. Essentially they

found that the priority effect depends on whether the first list (A-B) is tested before presenting the second list. They suggested that not testing the first list leads to negative transfer and a general rather than specific pair-wise interference (as indicated through comparison of within-subject and between-subject designs).

An intuitive approach to explaining these results might appeal to differences in the subjects' strategies when the first list is tested immediately as compared to when it is not. Given paired associate instructions simply to remember the pairs for a late memory test, an immediate test on A-B may suggest to the subject that s/he need no longer concentrate on remembering that list as the "later memory test" has been completed. In the case in which no test of the A-B pairs precedes presentation of the second list, the subject is likely to feel that it is important to retain the A-B information for the "later memory test" and expends effort during second list learning to do so. This will interfere with learning the new information and the extent of the interference will be largely determined by the degree of overlap between the lists.

In the standard Tulving recognition failure paradigm it is reasonable to suggest that this latter situation applies, i.e., the person must retain the list information in spite of the intervening word association task. In the FAI condition of experiment 1, a situation extremely similar to that described in Tulving and Watkins (1974) exists except

that the subject must create the interfering C response in addition to the other interfering aspects of the task.

It seems reasonable to expect that by replacing the word association task with a second list and not testing the first list prior to second list presentation, one creates a paradigm analogous to both the Tulving recognition failure paradigm and that described in Tulving and Watkins (1974). It would seem likely that if one expects that either the first term or the second term of the pair may be given as a cue on the memory task, that presentation of either A-C or D-B pairs as part of the second list would lead to (more or less) equivalent levels of interference in learning the second list but that the interference at the time of retrieval would depend on the interaction between the type of retrieval task and intervening list characteristics. As far as first list retrieval is concerned, the particular pattern of results predicted for the various groups and pair types does not differ from that predicted by Bartling and Thompson for experiment 1. The Tulving and Watkins (1974) results and explanations thereof merely confirm the necessity of specifying the relative importance of the target and intervening list such that both lists are attended to and thus the level of interference inherent in the A-B;A-C and A-B;D-B paradigms is also specified.

In order to assure that both lists were attended to in the second experiment, instructions at the beginning of the experiment included the fact that they would see two lists

of pairs of related words, that some words might appear in both lists, and that at some later point they would be asked to remember one list or the other. They were told that they would not be informed prior to presentation of the two lists as to which list would be tested. After the presentation of the lists and math filler tasks, the recognition and cued recall tasks of the previous experiment were performed always testing the first list. This procedure, then, minimizes differences in task difficulty between the two interference conditions not due to A-C;D-B differences and assures that the competing associations were encoded.

The issue of subject-generated versus experimenter-generated interfering associations cannot be completely resolved on the basis of the existing data. The Tulving and Watkins (1974) study does not include the requisite multiple retrieval situations to initiate comparisons. When Hunt (1975) used an experimenter-generated set of altered contexts, recognition maintained its high rate. However, the context change was introduced at recognition rather than prior to testing and subjects may have ignored the new cue as extraneous material. The instructions in the current experiment to the effect that either list may be tested make ignoring the intervening list impractical and should reduce the likelihood of this potential problem.

It is also possible that recognition in the BAI condition could be enhanced by the inclusion of the B terms

in both the original and intervening lists. On the other hand, this could create something of a list discrimination problem such that the person remembers seeing the word but cannot recall whether it was in the appropriate (first) or inappropriate (second) list, leading to conservative responding. However, both of these potential problems exist to some extent in the standard recognition failure paradigm. To alleviate the potential problem of facilitation of B recognition in the BAI condition (but not the FAI condition) due to B appearing in both the original and intervening lists, a number of untested E-F buffer items were included in both the BAI and FAI conditions of the second experiment. In this way the fact that all of the response terms appeared in both lists for the BAI group should be camouflaged and subjects should be less likely to be able to perform in the recognition task solely on the basis of having seen the responses at some indeterminate point in the experiment.

As the predictions of Bartling and Thompson's (1977) model which were stated in conjunction with the first experiment still hold and as the present paradigm was designed to test the same assumptions and implications of the model as the first experiment, these predictions will not be repeated here. It should be added, however, that the replication of the recognition failure phenomenon in the present BAI condition would not only add weight to the generalizability of recognition failure but may allow investigation of some previously unexamined features of the

paradigm and phenomenon.

A. Method

Subjects

Subjects for this experiment were 60 undergraduates at the University of Alberta who volunteered to participate. Of these, 56 were taking an introductory psychology course during the Spring session and received course credit for their participation and 4 had previously volunteered their services as subjects at the end of a Winter session psychology course. These subjects were randomly assigned to the experimental or control conditions on the basis of order of appearance at the lab such that there were no more than five subjects in a row assigned to a single group and there was a total of 20 subjects in each group. Subjects were run individually or in groups of between two and four in order to maximize the probability that instructions were understood and followed.

Materials

A practice trial identical in procedure to the critical trial used paired associate lists which were each composed of six word pairs related within the pair but unrelated to other list pairs or critical list pairs. The critical target list was composed of 10 of the 12 N-A pairs, 10 of the 12 A-N pairs, and 3 of the 4 primacy buffer items used in experiment 1. Two of each pair type from the previous experiment were dropped which had contributed the most

intrusions in the FAI condition in the previous experiment. (These items were Dark-ALLEY, Yellow-CAB, Cabbage-ROUND, and Home-SWEET.) An additional 5 untested buffer pairs were scattered throughout the critical list to bring the total list length to 28 pairs.

The BAI intervening list was composed of the corresponding 20 stimuli from the experiment 1 word association task paired with the appropriate original list responses with an additional 8 pairs unrelated to the original list pairs, 3 at the beginning of the list and 5 scattered through the list at irregular intervals. The FAI intervening list was composed of the same unrelated pairs plus the 20 stimuli from the original list paired with their most frequent response according to either Palermo and Jenkins (1964) or Keppel and Strand (1970) (or failing those sources, intuition aided by the experiment 1 protocols). The Control intervening list was composed of the 8 unrelated pairs in the previously mentioned lists plus an additional 20 pairs unrelated to those in the original list.

The critical list pairs were typed on a slide with the input cue having only its initial letter capitalized typed above the target which was completely capitalized. A test booklet consisting of 14 pages was provided for each subject, 7 pages dealing with each of the two sets of lists (practice set and critical set). The first page of each half contained instructions regarding the nature of the task. The second and fourth pages contained math problems.

The third page was blank. The fifth contained a 4-alternative free choice recognition task preceded by instructions. The sixth page was a cued recall of half of each type of pair input cues given their targets and the seventh page was a cued recall of the targets given their input cues for the other half of the pairs. With respect to the cued recall tasks, half the items were tested in the first cued recall for half the subjects and in the second cued recall for the other half of the subjects.

Procedure

A copy of the test booklet and a pencil was provided and the subjects were invited to read the instructions along with the experimenter. The instructions indicated that they would see two lists of word pairs, some words of which might appear in both lists. Following this, the experimenter would test them on one of the two lists, but they would not be informed as to which until test time. They were told that it was important not just to remember the words but also which word was paired with which, since they might, among other things, be given the first word and asked for the second or vice versa.

The task sequence was as follows with instructions at appropriate points: a) presentation of the first practice list at a 3-sec. rate via a Kodak Carousel projector, b) a 1 minute math computation task, c) presentation of the practice intervening list at the same rate, d) a 1 minute math computation task, e) a 1 minute recognition task, f) a

1 minute cued recall task for the input cues of half the pairs, and g) a 1 minute cued recall task for the targets of the other half of the pairs. The critical list presentation and testing followed the same procedure except that 3 minutes were allowed for the recognition task and 2 minutes for each of the cued recalls.

B. Results and Discussion

Noun-Adjective Differences

As in the previous experiment, the noun proved to be the more effective cue. The means presented in Table 4 show this effect in general for all three groups. However, as the control group has fewer additional factors complicating matters, the noun superiority effect will be examined in more detail using the control group results.

As before, when the noun serves as cue, as in the case of cued recall of the target for N-A pairs and recognition and cued recall of the input cue for A-N pairs, the mean percent correct is higher than for the corresponding cases in which the adjective serves as cue. As well, recognition failure was greater for the case in which the adjective is to be recognized while the corresponding noun serves as recall cue than vice versa.

The retrieval asymmetry ratios (RAR), as measured by #CRt/(#CRt + #CRc), for the Control condition demonstrate very clearly the superiority of the noun as cue for the N-A pairs (.76) but seemingly less clearly for the A-N pairs

Table 4

Mean Percent of Items Remembered, Recognition Failure, Recall Failure, and Retrieval Asymmetry Ratios

		Rn	CRc	CRt	$\%(\overline{Rn}\mid CRt)$	$\%(\overline{CRt}\mid Rn)$	RAR
A-N	Control	54.0	49.0	50.0	38.9	32.0	.49
	BAI	68.0	35.0	39.0	8.5	50.1	.52
	FAI	51.0	58.0	42.0	18.9	36.9	.43
N-A	Control	38.5	21.0	65.0	50.0	17.1	.76
	BAI	45.5	27.0	42.0	21.3	47.7	.56
	FAI	48.0	47.0	68.0	38.5	36.6	.49

Rn=%Correct Recognition

CRc=%Correct Cued Recall of the Input Cue Given the Target

CRt=%Correct Cued Recall of the Target Given the Input Cue

$\%(\overline{Rn}\mid CRt)$=%Recognition Failure of Recallable Targets

$\%(\overline{CRt}\mid Rn)$=%Recall Failure of Recognizable Targets

RAR=Retrieval Asymmetry Ratio=#CRt/(#CRt + #CRc)

(.49). This shift toward larger RAR values compared with the experiment 1 results appears to be largely due to an increased tendency for forward recall (CRt) to be superior to backward recall (CRc). While in experiment 1 this was a trend only, in experiment 2 this is more than a trend. In the main analysis of variance performed on only the three direct measures of Rn, CRc, and CRt for the three groups and both pair types (presented in Table 5), this superiority of forward over backward recall shows up as a significant effect for Measure with more or less equivalent levels of Rn and CRt and a lower level of remembering for CRc in all but one instance (FAI: A-N).

One would expect that if the directionality of the cuing had little or no effect, retrieval should be relatively asymmetric and lower than .50 for A-N pairs in the control condition. However, if forward cuing, as well as the noun as cue, produces superior recall, then these two opposing asymmetric trends would balance out producing an RAR value close to .50. Since for A-N pairs the naturally stronger cue acts in the weaker cuing situation and vice versa, the RAR can be reasonably expected to approach .50. In the same way the RAR value for the N-A pairs is higher than might be expected if directionality of cuing has no effect. However, the stronger cue acts in the stronger cuing situation in the N-A case, incrementing the already high RAR value.

This finding was statistically confirmed by a simple

Table 5

Main Analysis of Variance Summary Table

for Experiment 2 Recognition and Cued Recall Measures

Source	SS	df	MS	F
Group	25.35	2	12.68	<1
S(Group)	1081.93	57	18.98	
Measure	83.72	2	41.86	9.60*
Measure X Group	160.03	4	40.01	9.17*
Measure X S(Group)	497.25	114	4.36	
Pair type	36.74	1	36.74	6.52*
Pair X Group	14.74	2	7.37	1.31
Pair X S(Group)	321.03	57	5.63	
Pair X Measure	126.34	2	63.17	20.64*
Pair X Measure X Group	37.81	4	9.45	3.09
Pair X Measure X S(Group)	348.85	114	3.06	

*$p < .01$

effects analysis performed within the Control condition on the two types of cued recall separately for the N-A and A-N pairs (found in Appendix D). No real difference was found for the A-N pairs ($\underline{F}(1,19)=.03$, MSerror=3.40) and a significant difference between cued recall types for N-A pairs was found ($\underline{F}(1,19)=14.24$) favoring the forward recall situation.

<u>Assumptions of the Bartling and Thompson Model</u>

The superiority of forward over backward cued recall becomes something of a complicating factor in examining the assumptions and predictions of the Bartling and Thompson model, which was the original intent of the second experiment.

<u>Independence of Directional Associations</u>

The first of these assumptions to be examined is the independence of forward and backward associations. The model suggests that in comparison with the Control condition a decrement in performance should be specific to the cued recall type for which there was interference. Thus one would predict an interaction between cued recall type and Control versus Interference condition. Two separate analyses of variance were performed in connection with this prediction. The first compared the Control and FAI groups on the measures of CRc and CRt (cue type) summed across both pair types. The second analysis compared the Control and BAI groups in the same manner. A significant interaction was obtained for the FAI versus Control analysis

($\underline{F}(1,38)=22.78$, MSerror=7.38) but not the BAI versus Control analysis ($\underline{F}(1,38)=3.95$, MSerror=8.56). (See Appendix D for complete summary table.) However, it should be noted that in the former situation the interaction is due largely to better backward recall for the FAI condition rather than poorer forward recall in comparison with the Control condition. It is not immediately clear why a condition that is intended to produce interference should, in comparison to a non-interference control condition, perform at a superior level on those measures that are unaffected by the interference. However, a general facilitation of the non-interfered with directional association also appears in the experiment 1 data and, due to the lack of such a control in Tulving's and in Bartling and Thompson's research, it is difficult to determine the generality of the effect. Support for the assumption that forward and backward associations are affected independently of one another demands not simply that manipulations intended to produce interference with a particular direction of association lead to poorer performance on measures requiring retrieval of that association. That manipulation must also have no effect, facilitative or otherwise, on the other directional association. Regardless of whether this apparent facilitation of the non-interfered with association is a result of random error or reflects true facilitation, the results do not support the independence of associations assumption.

<u>Involvement of Directional Associations in Retrieval</u>

Because of various confounding effects in various conditions, it was difficult to determine how, if at all, directional associations were involved in retrieval. It was originally expected that on the recognition measure there would be a significant decrease for BAI in comparison with the Control condition but no significant decrease for FAI in comparison with the Control condition. Separate analyses of variance were performed involving only the recognition measure summing over pair type. One analysis compared the recognition performance of the BAI with the Control group. The other compared the recognition performance of the FAI with the Control group. While the latter result was obtained (see Appendix D), there was actually some slight improvement in recognition for the BAI condition ($F(1,38)=4.24$, $MSerror=10.91$). The probable cause of this improvement will be discussed shortly in relation to Bartling and Thompson's suggestion that recognition may be thought of as an implicit cued recall of the input cue. The statistical consequences of predicting and accepting a null hypothesis as in the case of the FAI-Control comparison are, of course, well known. All that can be stated is that the results in this case do not provide evidence to the contrary.

While the recognition failure measure ($\%(\overline{Rn}|CRt)$) in the previous experiment could be used at least in part to address the issue of directional association involvement in

retrieval, the Control condition in the present experiment was intended to serve as a standard control for interference and does not serve as a midpoint for the recognition failure measure largely due to the inflation of the BAI recognition rate and the strong forward cuing superiority. For these reasons the recognition failure data will not be discussed in relation to this issue.

Recognition as an Implicit Cued Recall of the Input Cue

The issue that is probably addressed most directly and conclusively by these data is the contention that recognition is essentially an implicit cued recall of the input cue. Differences between Rn and CRc can be accounted for by differences in the recognition-decision stage of remembering for the two tasks overall and, therefore, do not reflect seriously on the viability of this contention. However, if the retrieval processes involved are the same, then the pattern of relative levels of Rn for the three groups should mirror those for CRc. In point of fact, there was a significant interaction between Group (Control vs. FAI vs. BAI) and these two measures (Rn vs. CRc) summed over pair types ($F(2,57)=10.05$, MSerror=8.39). A simple effects analysis examining each group separately pinpointed more clearly the source of this interaction. While FAI showed almost no difference between these two measures ($F<1$), the Control condition produced somewhat better Rn than CRc ($F(1,57)=6.03$, $p<.017$) and BAI produced significantly better Rn than CRc ($F(1,57)=32.21$). (See Appendix E for complete

analysis of variance and simple effects summary.) The key to this interaction is the large difference in the BAI condition.

It was suggested in the introduction to this experiment that a potential problem existed in the BAI condition in that, unlike the other two conditions, the target appeared in both lists and a recognition decision might be possible on the basis of experimental familiarity. It was suggested that the inclusion of additional untested control pairs might camouflage the fact that the recognition task might be performed without regard to list membership problems. The results as obtained, however, indicate that this tactic probably had little or no effect in eliminating or even reducing the problem. Instead, it is likely that the recognition task was performed on the basis of experimental familiarity without the need to perform any list discrimination. The inflated recognition measure would, of course, be likely to result in a depressed level of recognition failure with respect to the Control condition and in general.

The CRc measure for BAI presents quite a different situation from Rn. While the target portion of the pair is present in both of the BAI condition lists, the input cue paired with the target is different in each list. The CRc task requires retrieval of the list 1 input cue for a correct response. While Rn can probably be performed on the basis of familiarity alone, CRc is heavily dependent on list

discrimination for successful performance.

The Role of List Discrimination in Remembering

The lack of dependence on list discrimination in the Rn task and the necessity of list discrimination in the CRc task demonstrates quite clearly that different processes may be involved in the two tasks or that at the least the retrieval-decision processes are involved to radically different extents. Recognition cannot be simply thought of as relying on backward associations as Bartling and Thompson (1977) have suggested, but may in some circumstances be relatively independent of the directional associations.

Recognition as a task often involves more than recognition as a process. Mandler (1980) has differentiated recognition and identification tasks, as well as presented a two-stage model of recognition bearing some similarity to Kintsch's (1970) model of recognition. Basically, Mandler suggests that a recognition task may be performed on the basis of experimental familiarity and failing that, a retrieval process will be invoked. This implies that very familiar items will not reach the retrieval stage and that there will be a positive decision solely on the basis of high situational familiarity. The retrieval process that Mandler (1980) proposes as his second stage is not inconsistent with the model of recognition proposed by Bartling and Thompson (1977). Nor is it inconsistent with the Tulving model in that the "retrieval processes involved in recognition are essentially the same as those used in

recall tasks" (Mandler, 1980, p. 256).

Retrieval Asymmetry and Recognition Failure

With respect to the suggested role of asymmetry in Tulving's recognition failure paradigm, the initial asymmetry in favor of the noun as cue was supported. The Control condition RARs demonstrate the strength of the noun superiority with the direction of recall factor taken into consideration. The analysis of variance performed on the Control group cued recall data (CRc and CRt) showed a significant Pair by Measure interaction ($F(1,19)=49.11$, MSerror=2.44), which is summarized in Appendix E.

The effect of the presence of a second list (as a substitute for a free association task) on retrieval asymmetry was confounded in terms of analyzing the recognition failure data because the BAI recognition task differed from the other recognition tasks in important ways. However, the analysis performed on the RARs for BAI versus FAI suggested that while relative levels of asymmetry predicted by the Bartling and Thompson model were obtained, the differences did not reach significance. The model suggests that BAI should maintain a higher RAR value than FAI across both pair types and that the N-A pair RARs should be higher than the A-N pair RARs across both conditions, but that there should be little or no interaction between condition and pair type. While all three F-values were non-significant (see Appendix F), the interaction $F(1,38)=.04$, MSerror=405.57 was much smaller than either the

Group $\underline{F}(1,38)=1.85$, MSerror=752.64 or the Pair type $\underline{F}(1,38)=1.52$. These data suggest that the use of an intervening list as opposed to a free association task is not a very effective means of affecting the basic level of retrieval asymmetry.

<u>Possible Task/Process Differences between Experiments 1 and 2</u>

One may only speculate at this point as to the reasons for this difference. It is possible that if Mandler's (1980) notions about the role of familiarity in recognition are correct, the familiarity increment differs sufficiently from producing the response to seeing the response to affect the need to retrieve the original trace with its concomitant associations. According to Mandler's (1980) model as specified, these data would imply that inclusion in an intervening list must increment situational familiarity more than producing that same word in a free association task since it is suggested here that a familiarity based decision is made in the former case but that a retrieval based decision is made in the latter case. However, a more reasonable approach is to claim that not all decisions to use familiarity versus retrieval in a recognition task are based on purely quantitative grounds. Rather, it is probable that the increment in situational familiarity is greater in the subject-generated case. However, the knowledge that the perceived high situational familiarity of the target in the recognition task may be due to the

generation of the item in the recent association task may either raise the threshold needed for recognition on the basis of familiarity or suggest to the person that the high familiarity is due solely to the free association task response. In the situation in which the target is experimenter-presented as part of two lists, there would be an increment in familiarity either without the increased suspiciousness as to its source or with the knowledge that the increase implies membership in both lists. Thus retrieval and list discrimination would be unnecessary. Such a model would be consistent also with the findings regarding the relative effects of recall trials and additional study trials on later memory trials (Slamecka & Graf, 1978; Thompson, Wenger, & Bartling, 1978; Wenger, Thompson, & Bartling, 1980).

As well as the source-of-familiarity explanation for this difference between the two paradigms, it is probable that the wrong trace is likely to be reinstated due to the recoding of the target during the association task. As such, the likelihood of recoding during the second exposure is greater for the subject-generated associations than for the experimenter-presented intervening list because the association task situation is not as similar to the original list presentation as is an intervening list situation. That is, the contextual cues and task demands are more discrepant from the original list presentation in the former case. However, while a recoding explanation is extremely useful

for explaining the evident list discrimination problem in the CRc task facing the BAI condition in both experiments 1 and 2 and the standard recognition failure paradigm as replicated in experiment 1, the obvious facilitation (as opposed to a reduced decrement that the recoding hypothesis projects) seems to require the addition presented in the source-of-familiarity explanation. In brief, while recoding of the target is a likely component of the standard recognition failure paradigm, the difference obtained here by replacing the association task with presentation of an intervening list requires more than recoding for an adequate explanation.

Evidence for the Role of List Discrimination

Probably the most important result that this experiment has to offer concerns the role of list discrimination as it relates to association retrieval in recognition and cued recall tasks. A case has already been made for the notion that a list discrimination need not be made and probably was not made in the case of recognition in the BAI condition. The case for the involvement of list discrimination as a component of the cued recall tasks has been incompletely presented to this point.

The evidence supporting the involvement of associative interference in the cued recall tasks comes not only from the relative levels of cued recall obtained by various groups but also from an analysis of second list intrusion errors in those groups. As expected, there were very few

second list intrusions for CRc in the FAI condition and for CRt in the BAI condition, those measures supposedly corresponding to the non-interference control group measures. As well, errors on the cued recalls in the Control group were minimal (.5%). Looking at intrusions in the Control protocols from the interference second lists (i.e., A-C and D-B pairs that were not seen by the Control group but which represent for the most part strong extralist associations), the second list intrusions accounted for 4% of responses from the BAI list in CRc protocols and 3.5% of responses from the FAI list in CRt protocols. On the measures which should show interference (CRt for FAI and CRc for BAI) second list intrusions accounted for 18% of responses in each case. The distribution of errors across pair types once again reflects the noun versus adjective by direction of cuing interaction. In the FAI condition there were 24 intrusions when the adjective was the cue and only 12 when the noun was the cue. In the BAI condition there were only 19 intrusions when the adjective was cue and 17 intrusions when the noun was cue, the balancing being due here to the backward, therefore weaker, nature of the cuing situation. An analysis of variance performed on the intrusion errors on the two cued recall measures of the BAI and FAI groups, summarized in Appendix F, demonstrated that this interaction was significant ($\underline{F}(1,38)=87.22$, $MS_{error}=.7711$). As the balancing of grammatical classes in the second list was nearly identical for the two

interference conditions, these data would also support the notion that list discrimination is a problem in the appropriate cued recall tasks.

Thus far it has been suggested that successful cued recall is dependent on retrieval of the appropriate missing half of the original episodic trace. If the retrieval cue has been learned as part of two distinctive traces with two distinctive targets (or input cues in the case of backward cuing), then a discrimination must be made between those two alternatives, provided they are both still available. If the two episodic traces are distinct mental representations of the original two presentations and if the retrieval cue is a copy of the overlapping element, then it is not clear whether the extracted retrieval information is more likely to contact one or the other of the two traces. The researcher must rely generally on the indirect evidence of the overt response in order to attempt to identify the characteristics of the retrieval cue as encoded. List discrimination as an explanatory concept depends on the retrieval cue accessing both traces. Thus the question of whether the traces are indeed separate and distinct needs to be addressed in order to determine the role that interference in general and list discrimination in particular must play in a successful model of episodic memory.

IV. The Separability of Episodic Traces: Experiment 3

In both Tulving's and Bartling and Thompson's models, events which are encoded after the target-event encoding and prior to retrieval should not interfere with the target's or input cue's effectiveness as a retrieval cue unless it in some way changes the features which are extracted from that target or input cue at retrieval (or in Bartling and Thompson's terms unless it sets up competing associations). In the standard recognition failure paradigm the target is directly presented as part of a different episodic trace in the free association task. So, if the input cue - target pair is "lamp-TABLE" and TABLE is then seen in the context of "chair", TABLE possibly might be less likely to overlap featurally with the original encoded trace because of the recoding. As mentioned before, however, the amount of featural overlap or recoding is not measurable separately from the amount of interference directly produced.

While both models do account for the effects of recoding part of the trace or directly setting up competing associations, the models also need to deal with the case of indirectly influential episodic experiences in which recoding of the cue should not be a factor. So, for instance, if "bubble-COOK" was presented in the critical list and "boil-HEAT" was presented later in the list, should one expect no particular decrement in the ability of "bubble" to lead to a response of "COOK"? While "bubble" and "boil" are similar in many respects including the way

each is used in the lists, is it reasonable to expect that the occurrence of one should change the meaning or extracted retrieval features of the other?

The relation between "bubble" and "boil" is one that has been traditionally referred to as an implicit associative response (IAR). In traditional S-R approaches and frequency theory the IAR was often called to mind during the encoding of a word. The more closely the word or concept was associated with a target, the more likely it was to be triggered as an IAR during encoding. One could specify which words were likely to be elicited as IARs to a particular target, according to the S-R approach, via which words were among the most common responses to that target in word association norms. In this wasy, a word which is generated internally by the subject is often treated as if it can be identified a priori by the experimenter.

Within a feature-encoding framework, the IAR would have a great number of features overlapping with those of the target. In addition, this type of feature model usually holds that some features are more important or heavily weighted than others. It therefore seems reasonable to assume that only a few of the important features would not overlap between a target and its IAR. It is clear that regardless of whether a more traditional S-R approach or a featural model is used, "bubble" in the example may be an IAR of "boil" in the example.

Because Bartling and Thompson's model conforms

generally to an association-based interference model, it seems reasonable to project that confusion would result if IARs of the input cues were paired with IARs of their respective targets. It is intuitively easy to see how one might become perplexed as to which of COOK or HEAT had been paired with "bubble" and which with "boil".

The meanings of bubble and boil are similar in general and become nearly identical in interaction with their targets. One would expect that the episodic traces in Tulving's sense should be unique. If the trace is available at all, the input cue should provide more retrieval information that matches the appropriate episodic trace than that matches an inappropriate trace. In cued recall the input cue should contact its own target as it has the most features in common with the episodic trace representation. Episodic traces are generally independent of one another in Tulving's model.

In the traditional recognition failure paradigm or even the type of paradigm used in experiment 2, a somewhat different situation exists from the one currently under discussion. One might view those paradigms as setting up situations in which the target is recoded during the free association task or during presentation of the second overlapping list. While the target is physically the same, the meaning is changed and the physical target is not necessarily a good retrieval cue for the original episodic trace. Mueller, Gautt, and Evans (1974) demonstrated the

necessity of reinstating the original encoding for cue efficacy. They used a modification of the A-B;A-Br transfer paradigm to test interference with list 1 associations. They presented double solution anagram-number pairs in either a standard A-Br or a modified A-Br situation. The critical modification was that one anagram solution was provided in parentheses for list 1 and the other solution provided for list 2. They tested retention of list 1 associations with a MMFR procedure given the anagram without any solution. Interference was reduced only if the original list 1 encoding could be reinstated, which might also be viewed as eliminating the list discrimination problem.

In the recognition failure paradigm, the critical list encoding of the target may not be reinstatable, or the recoded target may reinstate the wrong trace via backward associations to the free association cue. These two possibilities may be said to represent the positions of Tulving and of Bartling and Thompson when stated within the same framework. One can then readily understand why the recognition failure paradigm itself cannot provide a critical test for the two models.

Tulving (1974) presented a study that while not providing an adequate test of this distinction between a non-reinstatable encoding and reinstating the wrong encoding, suggested that such a test is possible and necessary. He modified the basic recognition failure paradigm by including a cued recall using the extralist cues

from the free association task prior to the free association-recognition and the cued recall using input list cues. Furthermore, the input cues were low associates of the targets but chosen to be congruous with the extralist cues in one condition and incongruous with the extralist cues in another condition. The general recognition failure result was in evidence, but cued recall with the congruous input cue was reduced with respect to the incongruous input list cue, significant at the .01 level. This finding, however, was not replicated when the first two retrieval tasks were reversed in order and, therefore, Tulving dismissed the first experiment result. It is possible that the congruous cues created some competing associations that the incongruous cues did not because of the recoding of the target. However, the necessity of such a recoding in one case and the lack of same in the other is certainly in question. In the previously presented experiment 2, successful performance would seem to demand that the two lists be separately and distinctly encoded and that recoding of overlapping elements not take place. Yet the data from intrusions suggested that some recoding did indeed take place. Also, the effect of order of task on the congruous/incongruous cue difference in Tulving's experiments is no more easily accounted for by this author given what is presented by Tulving than by Tulving himself. The lack of standardization of his cues either pre-experimentally or <u>post hoc</u>, and the order of task

effects which are left as epiphenomenal leave one feeling that the research conclusions are more to be stated in the form of questions than as declaratives.

Because presenting an item in two different contexts and then testing with that item does not allow control of the encoding used at test or the amount of recoding performed at the second encoding, the use of projected or normative IARs may provide an instance in which the uniqueness or non-uniqueness of the two episodic encodings is controlled to a greater extent and the word used at test is nominally and functionally identifiable.

Assume that pairs of associated words such as "bubble-COOK" were presented in a list such that for every pair presented earlier in the list another pair corresponding to it existed later in the list. This other pair could be an "IAR" pair such as "boil-HEAT" in which each element of the second pair is an assumed IAR of the corresponding element of the first pair. The encoding of the pair should serve to maintain this similarity of meaning relationship or encoding.

Alternatively, the second pair could contain a term which would be expected to be an IAR of an element of the first pair under one encoding, but is required by its paired term to use a different non-IAR encoding. An example of this would be "boil-SORE". Here a different meaning of boil is intended and one would expect the non-IAR encoding to predominate.

A third alternative could involve an arbitrary intrapair association but which contains the nominal IAR, such as "boil-CHEESE". Because the more common meaning of boil is an assumed IAR of bubble in its list usage, i.e., a major characteristic of boiling liquids is the production of bubbles, the encoding as an IAR should not be uncommon. However, both the possibility of some recoding and the lack of a firm semantic memory basis for the within-pair association would suggest that this pair's effect on the earlier encoding may set a boundary condition on the utility of the IAR concept in the description of the course of the episodic trace. It may be the case that the "unrelated" response term may result in an episodic encoding that is sufficiently distinct from the original critical pair encoding such that no interference such as that expected with IAR pairs like "boil-HEAT" and "bubble-COOK" would result.

Finally, a standard C-D pair such as "mouse-CHEESE" which maintains the intrapair relationship but is unrelated to the first pair provides a control condition. When retrieval of the encoded first pair is compared across the four conditions, the effect of the various types of relationships defined by the particular second pair encodings can be measured. The obtained differences add information which will certainly define the extent to which there is interaction between the episodic traces and may indicate the appropriate direction that either model must

take to remain viable.

It is unlikely that either a model which relies on meaning but does not take full account of association-based interference or a model which is association-based but does not provide any boundary conditions of association can be adequate. However, determining the stability and inviolability of the episodic trace and its elements may suggest which model provides a better base for an adequate description of episodic encoding and retrieval processes.

An adequate description of Tulving's model as stated in Tulving (1976) and Flexser and Tulving (1978) might suggest in this instance that the IAR relationship should have no effect so long as the target is distinguishable from the related word presented in the list. The input cue should not be reduced in its effectiveness in producing the target, in this example COOK. However, it is likely that intrusions should be in favor of HEAT as there is a good deal of featural overlap plus high situational familiarity. In brief, there should be no decrement in correct recall compared with the C-D control group (because there is no trace interaction) but intrusions should favor the associated response as decay of episodic trace features may lead to an increased similarity aided by the fact that it was also a list target.

The associative interference model would predict a decrement here because the presence of IAR pairs would create response competition (an intralist version of list

discrimination) and/or cue confusion. The presence of the related (IAR) pair may disrupt either or both forward and backward associations by presenting logical competing responses and cues which are similar and semantically related to each of the responses, albeit episodically distinct in the Tulving sense.

Perhaps the more interesting, but not critical to comparison of the models, are the conditions in which there is a word normatively predicted to be an IAR whose encoding is biased toward the non-IAR meaning and an assumed IAR paired with an unrelated word.

There is some indirect evidence concerning the effect of cuing two different meanings using homographs in a verbal discrimination task. Kausler (1973) presented homographs as correct items on one presentation and as wrong items on the other presentation. The non-homograph member of the pairs could either be associated to different meanings, both to the primary meaning, both to the secondary meaning, or unassociated with either meaning. The number of errors made over trials was calculated. It was suggested that when different meanings were cued on the two presentations (e.g. <u>net</u>-tennis; net-<u>income</u>) that the different associative contexts would "tag" different meanings and the items would behave as if unrelated, thus reducing errors. The uncued group (e.g. <u>net</u>-rest; net-<u>cheat</u>) produced more errors, overall not significantly different from the group for which the primary meaning was cued on both presentations,

suggesting that the subject does not attempt to differentiate the homographs automatically and "naturally" ascribes the primary interpretation to them. The most errors were produced when the secondary meaning was cued on both presentations.

The suggestion for the present study is that when the meaning of the nominal IAR that is cued is the non-IAR meaning, that pair will be treated as related and on the whole the condition will function similar to the present control condition. The condition in which the nominal IAR is presented with an unrelated item is analogous to Kausler's uncued control group and may be expected to behave similarly. The dominant meaning, which is also the functional IAR of a previously presented item, should be the one encoded and, therefore, some interference is expected according to the associative model. There is a question of degree of interference in that the first elements of the pairs are related and the second elements of the pairs are unrelated. The degree of separability of these events in terms of processing treatment is something which has not (to this author's knowledge) been previously tested.

The critical items for testing the notion of separability of related episodic events are, of course, those items which do not vary from list to list (e.g. bubble-COOK). Any differences across conditions connected with retrieval of these critical non-varying items must be due to their interaction with the items which do

vary from list to list. The homograph and control pairs for the various lists cannot be compared directly because the item interaction factor is confounded with the natural differences in retrievability which exist between, for instance, unrelated pair members (boil-CHEESE) and related pair members (mouse-CHEESE). The homograph and control pairs are only important to this examination in their projected possible effect on the critical pairs to which they are yoked.

In addition, the degree to which associatively related items may be considered analogous to homographic events is a matter of speculation. The critical study as described earlier using non-homographic pairs as the critical items for retrieval (e.g. bubble-COOK) would not pose serious problems as to the cue encoding at retrieval. However, it would not be clear, either in the present study or in a study in which the pairs in a single list contained homographs encoding both meanings, what the retrieval cue encoding of a homograph was in any particular instance. As such, on the cued recall trial it is questionable as to which encoding of "boil" would occur, since both encodings are equally likely episodically albeit not semantically. Since it is not possible to predict which encoding would be reinstated, it is impossible to use solely homographic input cues to test or extend either model.

The associated stimulus terms, such as bubble and boil, allow more control of the encoding at test time and thereby

allow one to compare the two models and assess a few of the boundary conditions that intrapair association and alternate encodings impose.

A. Method

Subjects

Subjects were 80 undergraduates at the University of Alberta who participated in order to fulfill an introductory psychology course requirement. These subjects were randomly assigned to the four conditions on the basis of order of appearance such that there were 20 subjects in each group. Subjects were run individually or in groups of two or three in order to maximize the probability that instructions were understood and followed.

Materials

The four critical lists were composed of a set of 32 pairs of moderate to high frequency words, 43-AA for homographs and 1-AA for all others (Thorndike & Lorge, 1944). Each list contained four primacy pairs, 14 pairs of related items which comprised half of all four lists, and 14 pairs of items which comprised the other half and which varied from list to list. For condition A, both members of the pair were probable IARs of a pair presented earlier and maintained the same basic relationship. In addition, the stimulus term was a homograph. For condition B, the stimulus terms were nominally IARs of the earlier stimulus terms but were paired with responses such that the

alternative meaning of the homograph was biased. All homographs used in conditions A and B, their response terms, and the related critical pairs presented earlier in the list were taken from the Cramer (1970) word association norms for homographs. All homographs had a dominant meaning related to the corresponding stimulus in the related critical pair in the list. For condition C, the stimulus terms were nominally IARs of earlier stimulus terms and were each paired with an unrelated response. The same homographs as in conditions A and B were used here as well. Finally, for condition D, neither member of the pair was a likely IAR of anything earlier in the list but the same type of intrapair relationship was maintained as in the critical pairs. The pairs used in the other half of list D were composed of the response terms from the unrelated pairs in the non-critical half of list C paired with associated stimulus terms. Table 6 presents the pairs from all four lists in experiment 3.

Each stimulus-response pair was typed on a slide with the stimulus term printed above the response term, both completely capitalized. In addition, each stimulus term was typed on a slide to be used in the test portion of the task. Each critical pair was presented an average of 6.1 items before its matched pair (SD=1.17).

Procedure

Subjects were given a written copy of the instructions and asked to read them along with the experimenter. A practice list of 14 pairs of related items was used to

Table 6

Stimulus-Response Pairs for Experiment 3

Critical Pairs		Matched Pairs		
	Condition A	Condition B	Condition C	Condition D
Bubble-COOK	Boil-HEAT	Boil-SORE	Boil-CHEESE	Mouse-CHEESE
Container-SQUARE	Box-CARDBOARD	Box-FIGHT	Box-JACKET	Leather-JACKET
Group-SOCIAL	Club-JOIN	Club-BEAT	Club-TIME	When-TIME
Law-JUDGE	Court-JURY	Court-TENNIS	Court-CLARINET	Oboe-CLARINET
Goose-FLY	Duck-SWIM	Duck-DODGE	Duck-PERFUME	Lotion-PERFUME
Cabinet-STEEL	File-CARD	File-NAIL	File-SPICE	Sugar-SPICE
Good-NICE	Fine-GREAT	Fine-PAY	Fine-AUTUMN	Leaf-AUTUMN
Thin-SKINNY	Lean-SLIM	Lean-AGAINST	Lean-CLOUD	Sky-CLOUD
Love-HUSBAND	Like-FRIEND	Like-SAME	Like-LIME	Lemon-LIME
Parade-WALK	March-STEP	March-MONTH	March-VALLEY	Hill-VALLEY
Book-STUDY	Page-READ	Page-CALL	Page-SPIDER	Insect-SPIDER
Harbor-SHIP	Port-BOATS	Port-WINE	Port-WISH	Dream-WISH
Win-HORSE	Race-CAR	Race-NEGRO	Race-SEAT	Chair-SEAT
Deny-ACCEPT	Refuse-OBEY	Refuse-WASTE	Refuse-COUNTRY	Patriot-COUNTRY

familiarize the subjects with the experimental procedure. The word pairs were presented at a 3-sec rate via a Kodak Carousel projector. A 1-min math computation task followed as a control for recency effects. Then instructions were given to recall the response terms given the appropriate stimulus terms. Each stimulus term was presented for 7 seconds via projector and the subject wrote down his/her response on a sheet of paper with numbered lines corresponding to each stimulus term. The same procedure was followed with the critical list.

B. Results and Discussion

While all pairs were tested, only the critical pairs for each condition were included in the statistical analyses. The means and standard deviations for correct responses on the critical and non-critical items are presented in Table 7.

There are no significant differences among the means for the critical items as would be expected if the traces were independent of one another. The non-critical items showed greater differences but as these items varied completely from list to list, no statistical analysis was deemed warranted. It may be noted, however, that the control list (D) items were the best recalled of the non-critical items and that they did not differ in recall from the critical items in the same list. Non-critical items in list C showed the poorest recall as might be

Table 7

Means and Standard Deviations for Correct Responses on Critical and Non-critical Pairs

	A	B	C	D
Critical Pairs				
Mean	9.15	8.60	8.90	9.35
SD	2.13	2.52	3.26	2.35
Non-Critical Pairs				
Mean	6.75	7.55	4.35	9.40
SD	2.88	3.09	2.48	3.05

expected given that they were unrelated within pairs. The Cramer (1970) norms gave equivalent association values for A and B list pairs. The recall means showed a slight advantage for the B list pairs which were composed of homograph stimuli biased toward their less frequent usage, .

An error analysis performed on the critical items for each group did indicate some differences. The number of errors made of various types for each condition appear in Table 8. As is evident from the table, far more within-list response errors were made in Condition A than in any of the other conditions and that this increase was accompanied by a corresponding decrease in no response and non-list responses. As well, the majority of within-list stimulus and response errors in condition A protocols were from the corresponding homograph pair.

Since the number of errors of a particular type is not independent of the number of errors of another type, the error distribution scores were transformed to reflect the number of subjects contributing to a particular error class. That is, each subject who made one or more errors of a particular type contributed a "1" to that class. Specifically, these error classes were 1) within-list stimulus, 2) within-list response, 3) practice list word, 4)non-list word, and 5) no response made. A series of Chi-Square analyses was then performed comparing groups for each error type calculated on this basis. The number of subjects from each group contributing to these error types

is presented in Table 9. Within-list response intrusions were made by more subjects in Condition A than in the other conditions (Chi-Square(3)=16.19). Within-list stimulus intrusions were made by only 1 subject in Condition D, while 13 out of 20 subjects in Condition A made errors of this type. Conditions B and C fell in between. Again, differences in number of within-list stimulus intrusions were substantial (Chi-Square(3)=15.76). While the numbers of "no response" errors may be considered to be markedly different across conditions, contribution to an error type did not reach significance (Chi-Square(3)=7.78, $p<.06$). This is largely due to the fact that nearly all subjects failed to respond to one or more of the critical stimuli.

The results are certainly consistent with those predicted by an extension of Tulving's (Tulving, 1976; Flexser & Tulving, 1978) model. It might be added that some people in condition A mentioned the difficulty of the task or concurred with the experimenter during debriefing that that particular task might be highly interfering. None of the participants in the other conditions made mention of the ease or difficulty of the task. This, however, should not be given much weight as it may be due at least in part to biasing by the experimenter's debriefing or may reflect differences in uncertainty levels generated by the list characteristics. It is possible that the lack of differences in the number of errors for critical items among the conditions is due simply to the person in the more

Table 8

Number of Errors on Critical Items of Each Type for Each Condition

	A	B	C	D
Within-list Response	20	6	4	3
Within-list Stimulus	16	9	11	1
Practice List	4	7	2	8
Non-list Response	22	36	29	37
No Response	35	50	56	44

Table 9

Number of Subjects Contributing Errors for Each Error Type and Condition

	A	B	C	D
Within-list Response	13	5	3	3
Within-list Stimulus	13	8	9	1
Practice List	4	6	2	7
Non-list Response	13	13	12	15
No Response	19	18	20	15

difficult situation applying more effort in order to perform as well as someone who had an easier list. This would explain why one might find the task difficult but still do as well as another with an easier task. However, a plausible alternative is that when a person in the projected easiest condition (D) did not know the answer, s/he either made a random (non-list) guess or made no response with little uncertainty as to its incorrectness and little effort. When a person in the projected most interfering condition (A) did not know the answer, s/he chose a list item with great uncertainty and much effort. It is probably impossible to determine how much of each of these alternatives contributes to these results. At this point, however, it may be claimed that these data provide tentative support for Tulving's model of remembering as stated in Tulving (1976) and in Flexser and Tulving (1978) and for the possibility of separable episodic memory traces when recoding is controlled.

V. General Discussion

It cannot be denied that retrieval asymmetry plays some part in the recognition failure phenomenon. In experiment 1, a very low level of recognition failure was demonstrated in the situation in which asymmetries naturally favored the recognition situation and no association task was employed that would reduce recognition in relation to recall. As demonstrated in the replication of the standard Tulving task, the natural superiority of the noun as retrieval cue is unquestionably supported.

However, retrieval asymmetry factors cannot account for the whole recognition failure phenomenon. Tulving and Thomson (1973) have already suggested that the notion of differential strength of forward and backward associations (and it may be implied that differential cue efficacy of nouns and adjectives would also be included) as an explanation for recognition failure is subject to the reverse interpretation. That is, the "backward association appears to be weaker than the forward association whenever the target word is not as readily recognized as the cue word" (Tulving & Thomson, 1973, p.366). The basic view of Bartling and Thompson (1977) and Rabinowitz et al. (1977) has been that the target fails to be recognized because it does not implicitly retrieve the input cue. Flexser and Tulving (1978) stated that none of the data to date had been able to distinguish between this interpretation and the alternative that retrieval of the cue given the target fails

because the target is not recognized.

One may suggest that the data from the second experiment distinguish between these two alternatives. Specifically, the BAI manipulation resulted in a high level of recognition but a low level of cued recall of the input cue given the target with respect to control group performance. The interpretation given to this finding was that retrieval of the specific first list association was interfered with but that the recognition task was likely being performed without the necessity of a backward retrieval. It would therefore become unlikely that failure of the implicit backward retrieval can be used to explain the recognition failure phenomenon since it is not a necessary component of recognition in some situations.

The limitations on the role of association retrieval as part of the recognition processes suggests that a reevaluation of the current models of recognition and recall are in order. It seems fairly clear that Bartling and Thompson's (1977) model of the two processes as each involving the use of appropriate directional associations is limited in its generality to situations in which some sort of trace discrimination (of which list discrimination is one sort) must be made. Flexser and Tulving's (1978) description of recognition and recall as each involving extraction of retrieval information and comparison with trace information, on the other hand, is so general as to be capable of handling both types of situations but not

specific enough in delineation of how retrieval information is extracted to distinguish in all cases when trace discrimination information is necessary to successful remembering.

The suggestion that recall and recognition involve the same basic underlying processes has been evident throughout this discussion and allows for an elegant model of memory. However, as demonstrated by Bartling and Thompson's and Tulving's models, one can become overly rigid or overly flexible in one's description of memory processes. The literature in this field is vast and while a flexible model such as Tulving's is at a distinct advantage in handling the abundance of memory phenomena already well-researched, one can only infer how the general model would deal with the specific phenomena on a <u>post</u> <u>hoc</u> basis. The research presented here suggests that a discussion of the roles of interference and associations in interference with respect to recall and recognition processes in Tulving's general model is in order.

As stated previously, recall and recognition both involve the extraction of retrieval information from the situational cues and matching of that information to that found in the episodic trace. Interference can be associated with one of several substages of remembering. For instance, one could extract the wrong features, i.e., features not in common with those in the episodic trace. This is largely due to recoding of the retrieval cue such that the

connotative meaning is changed (Light & Carter-Sobell, 1970; Reder, Anderson, & Bjork, 1974). It has been suggested that in the standard recognition failure paradigm recognition may fail because the copy cue is interpreted in relation to the word to which it was generated, that is, it has been recoded in terms of the free association task context. In cued recall situations this might result in the overt production of a word related to the cue but unrelated to the target or in no response to the cue.

Another possible source of interference within Tulving's model is that extracted features could be compared to the wrong but similar episodic trace. (It should be noted here that some interpretations of Tulving's model would not allow this possibility due to the correct episodic trace always being chosen for the feature matching process.) This explanation includes the assumption that episodic traces are not always unique and separable and may, in fact, involve some recoding. Tulving (1976) has suggested that during encoding one may not only incorporate elements of the semantic memory into the trace but also elements of previous episodic traces. Thus, it is possible that in experiment 2 the overlapping elements in the two lists may have resulted in a second list encoding that included a "first list tag", or even a retrieval of the first list encoding and some inclusion of a "second list tag" in a recoding of it. Whatever the actual mental procedure, the end result might be the elimination of the need for trace discrimination in a

recognition situation such as that found in the BAI condition of experiment 2 and an increase in interference in a cued recall situation in which the overlapping element serves as cue. As a result, in this type of cued recall situation one would expect a large number of intrusions from the similar episodic traces.

A final possibility for interference is at the post-ecphoric decision stage which has, until now, not been mentioned because the models being compared largely deal with the more automatic or somewhat less consciously controlled retrieval processes (which Tulving (1976) has labelled "ecphoric" processes). Tulving (1976) suggested that "the subject may decide <u>after</u> he becomes consciously aware of the product of retrieval, that what he remembers does not fit what he wants" (p.65). However, this discrepancy is supposedly due to inappropriate or inadequate initial ecphoric information (retrieval cues). The alternative of a conscious override of the retrieved product is intuitively valid. In a standard recognition failure paradigm, recognition may possibly fail because the person judges the retrieved items as very familiar solely due to the fact that s/he explicitly or implicitly produced them in the previous association task. In the case in which recognition is performed on experimenter-presented items, recognition may fail because the person had produced a similar (overlapping) set of responses in the previous task and attributes the strong familiarity to the previous task.

The difference between these situations is in degree of response alternative overlap but not necessarily the presence or absence of situational familiarity accruing to the targets in the recognition task. Thus, changes in the number and similarity of alternatives in the recognition task would produce somewhat less than major changes in the level of recognition failure produced, but the absence of the association task would produce a much larger decrement in recognition failure. The difference in effect between the experiment 1 and experiment 2 results could in fact be as much related to the post-retrieval decision processes as to recoding or the lack thereof. It has been suggested in the earlier discussion of the experiment 2 results that the effect produced in the BAI condition by using a second list rather than an association task to produce interference might be explained in terms of a difference in the attribution of the increase in familiarity due to a second exposure to the target.

On the whole, decision interference, as opposed to retrieval interference, is problematic as an explanation of what may happen in remembering because it relies so heavily on subject strategies. The present studies do not distinguish between the possible sources of interference nor is it necessary to do so at this point. Rather, it is suggested here that all three sources of interference are possible within Tulving's model and are suggested in a _post hoc_ fashion by the particular patterns of results obtained

in the present experiments and others in the literature. Thus, a single model of recognition and recall processes based on Tulving's general model is possible when the conditions which may produce interference such as that found in the recognition failure paradigm and its variations are elucidated.

The role of associations in this interference is dependent on the sort of associations to which one refers. In terms of the forward and backward associations discussed by Bartling and Thompson (1977), it is clear that these associations are sometimes but not always a part of retrieval. As demonstrated in experiment 2, recognition in the BAI condition could not have required the use of backward associations or else the level of recognition performance would have mirrored that of the backward cued recall. However, it is also clear from these same data that backward and forward associations may and sometimes must play a role in some retrieval situations such as the two cued recall tasks in the first two experiments. As to associations' role in interference, these results would suggest that a competing association may interfere by affecting which features are extracted from the retrieval information (a recoding problem) or by leading to a match with the wrong episodic trace, both of which are problems in the retrieval stage.

It should be noted that Tulving's general model does not specify how features are extracted from the retrieval

cues. Clearly, people make use of very basic information such as the knowledge of the nature of the task, e.g. recall or recognition, in retrieval and it may be the case that the use of forward and backward associations <u>where appropriate</u> is simply a part of this feature extraction process. In this way the two general models being discussed are not incompatible, but Bartling and Thompson have merely been more specific as to the mechanisms used in retrieval. As such, forward and backward associations may play a part in feature extraction when task demands make them appropriate.

In terms of IAR-type associations, these play an important role in determining the types of features that are extracted at retrieval time as well as affecting the degree of recoding (or accession of the original trace) during storage of the second list or during production of the target in the free association task. One of the major difficulties in studying retrieval processes with the recognition failure paradigm has been the lack of control over the encoding of the retrieval cue. However, this problem seems to have more bearing on the nature of the episodic encoding than on recall - recognition process differences.

An episodic encoding has been defined by Tulving (1976) as an "item-in-context" which makes use of semantic memory information, previous related episodic memory experiences, and information about the current context in the encoding process. This definition allows for a certain overlap

between traces via tapping into common semantic memory or episodic memory information but still allows for a certain level of uniqueness and separability via independent selection of features for encoding (and forgetting) and the non-overlapping elements of the contexts.

Retrieval interference between two traces, then, should depend on two conjoint factors: a large amount of featural overlap during encoding and the inability of the specific features extracted from the retrieval situation to distinguish between the two traces. If there is little featural overlap (similarity) or if the extracted features of the retrieval cue more closely correspond to one rather than the other trace, then there will be little or no retrieval interference.

In experiments 1 and 2 the extracted retrieval information might correspond most closely to either the critical list presentation or the second presentation with little possibility for experimenter control of the encoding. Experiment 3 attempted to control the retrieval encoding so that it would more closely correspond to the critical item but vary the amount of featural overlap between the target and corresponding pair. As one might expect, if this notion of retrieval interference holds, there was no increase in interference on the critical items as featural overlap increased. As well, large amounts of featural overlap led to more guessing of the familiar, related list item than when there was little featural overlap.

This two-factor description of retrieval interference is somewhat different from that presented by Runquist (1975). He talks about "trace interaction" interference as involving stimulus coding errors and/or competing associations. The former, also labelled cue confusion, may be thought of as an inability to discriminate among the retrieval cues or a miscoding of the retrieval cue. The competing associations notion is functionally equivalent to the featural overlap of the encoded episodic traces discussed previously. Runquist has suggested that elimination of cue confusion via discriminative coding is neither necessary nor sufficient for elimination of interference. Unlike the present suggestion that both featural overlap during encoding and inability of the extracted retrieval features to discriminate among traces are necessary for interference, Runquist would hold that elimination of one factor does not necessarily eliminate the interference. This discrepancy is due to a definitional discrepancy between Runquist's cue confusion and the present extracted feature retrieval failure.

Runquist (1975) looks at cue confusion in terms of the ability to discriminate one cue from another without the additional aspect of retrieval efficacy and tests it via a method which does not demand original trace retrieval. The missing item recognition (MIR) test used involves presenting all but one of the stimulus terms in a standard paired-associate test sequence and then presenting either

the missing stimulus or one of the already tested stimuli for recognition as the remaining untested stimulus (missing) or as previously tested (repeat). In order to perform the recognition task, one must be able to discriminate that stimulus from the others in the experiment, but not necessarily be able to contact the appropriate response term. Runquist (1975) presents data showing that the level of interference in retrieving the appropriate response may decrease without a concurrent decrease in cue confusion. Also reported in Runquist (1975) were various experiments in which cue confusion was virtually eliminated without significantly affecting retrieval interference. Given that the MIR task does not demand original trace retrieval, it is not surprising that performance on it may function independently of performance on a task that demands trace retrieval (recall). Such a result was also obtained in the present experiment 2.

It is therefore suggested that while on the surface the present hypotheses as to the factors operating in retrieval interference appear to conflict with those of Runquist (1975), the two models are addressing different issues with respect to cue confusion.

In summary, it is proposed that a description of recall and recognition retrieval processes that depends on using forward and backward associations between cue and target is insufficient to deal with all episodic memory situations due to the fact that trace discrimination is not always a

necessary component obviating the need for full trace retrieval. Tulving's flexible model is more in keeping with the existing data and the interpretations of Mandler (1980) and those presented here. The Bartling and Thompson directional associations are almost certain to be part of the extracted retrieval information in some cases, but the general model does not require them. In addition, Tulving's model has been adapted to account for interference phenomena and to suggest in what respects episodic traces may be considered as unique and how they may interact.

On the whole, one may claim that recall and recognition processes may be described in terms of the same general information extracting, matching, and decision processes, although what information is extracted, what it is matched to, and the basis for the final decision may vary from one situation to another. It is hoped that the current discussion has provided useful expansions of the general model which will elucidate to some extent the factors controlling these variable quantities.

References

Anderson, J. R., & Bower, G. Recognition and retrieval processes in free recall. _Psychological Review_, 1972, _79_, 97-123.

Bahrick, H. Two-phase model for prompted recall. _Psychological Review_, 1970, _77_, 215-222.

Bartling, C. A., & Thompson, C. Encoding specificity: Retrieval asymmetry in the recognition failure paradigm. _Journal of Experimental Psychology: Human Learning and Memory_, 1977, _3_, 690-700.

Bowyer, P. A., & Humphreys, M. S. Effect of a recognition test on a subsequent cued-recall test. _Journal of Experimental Psychology: Human Learning and Memory_, 1979, _5_, 348-359.

Cramer, P. A study of homographs. In L. Postman & G. Keppel (Eds.), _Norms of word association_. New York: Academic Press, 1970.

Flexser, A., & Tulving, E. Retrieval independence in recognition and recall. _Psychological Review_, 1978, _85_, 153-171.

Horowitz, L., & Prytulak, L. Redintegrative memory. _Psychological Review_, 1969, _76_, 519-531.

Hunt, R. R. How similar are context effects in recognition and recall? _Journal of Experimental Psychology: Human Learning and Memory_, 1975, _1_, 530-537.

Jenkins, J. J. The 1952 Minnesota word association norms. In L. Postman & G. Keppel (Eds.), _Norms of word_

association. New York: Academic Press, 1970.

Jones, G. V. Recognition failure and dual mechanisms in recall. Psychological Review, 1978, 85, 464-469.

Kausler, D. H. Homographs as double function items in verbal discrimination learning. Journal of Experimental Psychology, 1973, 98, 214-216.

Keppel, G., & Strand, B. Free association responses to the primary responses and other responses selected from the Palermo-Jenkins norms. In L. Postman & G. Keppel (Eds.), Norms of word association. New York: Academic Press, 1970.

Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), Models of human memory. New York: Academic Press, 1970.

Kintsch, W. The representation of meaning in memory. Hillside, N. J.: Laurence Erlbaum & Associates, 1976.

Kintsch, W. More on recognition failure of recallable words: Implications for generation-recognition models. Psychological Review, 1978, 85, 470-473.

Light, L., & Carter-Sobell, L. Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 1-11.

Mandler, G. Recognizing: The judgment of previous occurrence. Psychological Review, 1980, 87, 252-271.

Mueller, J., Gautt, P., & Evans, J. H. Stimulus encoding in A-Br transfer. Journal of Experimental Psychology, 1974, 102, 54-61.

Palermo, D., & Jenkins, J. J. Word association norms. Minneapolis: University of Minnesota Press, 1964.

Rabinowitz, J. C., Mandler, G., & Barsalou, L. W. Recognition failure: Another case of retrieval failure. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 639-663.

Reder, L., Anderson, J. R., & Bjork, R. A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 1974, 102, 648-656.

Runquist, W. Interference among memory traces. Memory and Cognition, 1975, 3, 143-159.

Salzberg, P. M. On the generality of encoding specificity. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 586-596.

Santa, J., & Lamwers, L. Encoding specificity: Fact or artifact? Journal of Verbal Learning and Verbal Behavior, 1974, 13, 412-423.

Slamecka, N., & Graf, P. The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 592-604.

Thompson, C., Wenger, S., & Bartling, C. How recall facilitates subsequent recall: A reappraisal. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 210-221.

Thorndike, E. L., & Lorge, I. The teacher's word book of 30,000 words. New York: Columbia University, 1944.

Tulving, E. Episodic and semantic memory. In E. Tulving &

W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Tulving, E. Recall and recognition of semantically encoded words. Journal of Experimental Psychology, 1974, 102, 778-787.

Tulving, E. Ecphoric processes in recall and recognition. In J. Brown (Ed.), Recall and recognition. London: John Wiley & Sons, 1976.

Tulving, E., & Thomson, D. Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973, 80, 352-373.

Tulving, E., & Watkins, M. J. On negative transfer: Effects of testing one list on the recall of another. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 181-193.

Tulving, E., & Wiseman, S. Relation between recognition and recognition failure of recallable words. Bulletin of the Psychonomic Society, 1975, 6, 79-82.

Wallace, W. P. Recognition failure of recallable words and recognizable words. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 441-452.

Watkins, M. J., & Tulving, E. Episodic memory: When recognition fails. Journal of Experimental Psychology: General, 1975, 1, 5-29.

Wenger, S., Thompson, C., & Bartling, C. Recall facilitates subsequent recognition. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 135-144.

Appendices A-F

Additional analyses of experiment 1 and 2 data

Appendix A

Summary of Simple Effects and Contrasts

Testing the Assumptions and Implications of

Bartling and Thompson's (1977) Model

Source	SS	df	MS	F
CR type X Cntrl vs FAI	1665.30	1	1665.30	1.39
Cntrl vs FAI at CRc	4906.23	1	4906.23	4.09
Cntrl vs FAI at CRt	152.10	1	152.10	<1
CR type X S(Group)	45530.08	38	1198.16	
CR type X Cntrl vs BAI	2726.10	1	2726.10	3.22
Cntrl vs BAI at CRc	336.40	1	336.40	<1
Cntrl vs BAI at CRt	3080.03	1	3080.03	3.64
CR type X S(Group)	32164.72	38	846.44	
Cntrl vs BAI at Rn	6786.03	1	6786.03	6.46
S(Group)	39915.95	38	1050.42	
Cntrl vs FAI at Rn	2073.60	1	2073.60	1.78
S(Group)	44150.30	38	1161.85	
CRc vs Rn X Group	6705.51	2	3352.76	2.95
CRc vs Rn at Cntrl	10660.23	1	10660.23	9.37*
CRc vs Rn at BAI	1537.60	1	1537.60	1.35
CRc vs Rn at FAI	152.10	1	152.10	<1
CRc vs Rn X S(Group)	30731.58	27	1138.21	

*$p < .01$

Appendix B

Analysis of Variance Summary for the Recognition Failure and Recall Failure Data

Source	SS	df	MS	F
Recognition Failure				
Group	8451.15	2	4225.58	97.68**
S(Group)	2465.85	57	43.26	
Pair type	26344.03	1	26344.03	16.99**
Pair type X Group	5548.85	2	2774.43	1.79
Pair type X S(Group)	88383.95	57	1550.6	
Total	131194.00	119		
Recall Failure				
Group	4416.62	2	2208.31	2.79
S(Group)	45122.47	57	791.62	
Pair type	8755.21	1	8755.21	7.52*
Pair type X Group	49.81	2	24.91	<1
Pair type X S(Group)	66343.48	57	1163.92	
Total	124687.59	119		

*p<.01

**p<.001

Appendix C

Summary of ANOVAs Related to the Initial Level and Association Task Effects on Retrieval Asymmetry

Source	SS	df	MS	F
Initial Asymmetry (Cntrl)				
Pair type	159.61	1	159.61	<1
CR type	2194.51	1	2194.51	3.86
Pair type X CR type	11257.50	1	11257.50	19.86*
Error(pooled)	43172.77	76	568.06	
Total	56784.39	79		
Retrieval Asymmetry Ratios				
Group(Cntrl vs BAI)	4545.11	1	4545.11	8.75*
S(Group)	19737.38	38	519.40	
Pair type	10283.11	1	10283.11	22.09**
Pair type X Group	485.11	1	485.11	1.04
Pair type X S(Group)	17685.78	38	465.42	
Total	52736.49	79		
Group(Cntrl vs FAI)	2.45	1	2.45	<1
S(Group)	12522.34	38	329.54	
Pair type	9680.00	1	9680.00	11.69*
Pair type X Group	361.25	1	361.25	<1
Pair type X S(Group)	31473.76	38	828.26	
Total	54039.80	79		

*p<.01 **p<.001

Appendix D

Summary of Simple Effects and Contrasts Analyzing Experiment 2 Data

Source	SS	df	MS	F
Group(Cntrl vs FAI)	7.20	1	7.20	<1
S(Group)	908.6	38	23.91	
CR type	51.20	1	51.20	6.93
CR type X Group	168.20	1	168.20	22.78*
CR type X S(Group)	280.60	38	7.38	
Group(Cntrl vs BAI)	88.20	1	88.20	2.38
S(Group)	1408.60	38	37.07	
CR type	204.80	1	204.80	23.92*
CR type X Group	33.80	1	33.80	3.95
CR type X S(Group)	325.40	38	8.56	
Cntrl vs BAI at Rn	46.26	1	46.26	4.24
S(Group)	414.58	38	10.91	
Cntrl vs FAI at Rn	4.23	1	4.23	<1
S(Group)	487.56	38	12.83	
CR type X Pair type(Cntrl)	23.12	1	23.12	6.80
CR type at N-A	48.40	1	48.40	14.24
CR type at A-N	.10	1	.10	.03
Error	64.57	19	3.40	

*p<.01

Appendix E

Summary of Simple Effects

for Experiment 2 Data (continued)

Source	SS	df	MS	F
Group	160.80	2	80.40	2.36
S(Group)	1945.7	57	34.14	
Measure(Rn vs CRc)	86.70	1	86.70	10.33*
Measure X Group	168.80	2	84.40	10.05*
Rn vs CRc at FAI	3.60	1	3.60	<1
Rn vs CRc at BAI	270.40	1	270.40	12.72*
Rn vs CRc at Cntrl	50.63	1	50.63	2.38
Measure X S(Group)	478.50	57	8.39	
Pooled error	2423.64	114	21.26	
CR type at Cntrl	76.05	1	76.05	21.11*
Error(within CR type)	68.45	19	3.60	
Pair type at Cntrl	14.45	1	14.45	5.08
Error(within Pair type)	54.05	19	2.84	
CR type X Pair type	120.05	1	120.05	49.11*
Error(within CR X Pair)	46.36	19	2.44	

*p<.01

Appendix F

Analysis of Variance Summaries for Retrieval Asymmetry Ratios and Second List Intrusion Errors

<u>Source</u>	<u>SS</u>	<u>df</u>	<u>MS</u>	<u>F</u>
<u>RAR Values</u>				
Group(FAI vs BAI)	1394.45	1	1394.45	1.85
S(Group)	28600.25	38	752.64	
Pair type	616.05	1	616.15	1.52
Group X Pair type	14.45	1	14.45	.04
S(Group) X Pair type	15411.55	38	405.57	
<u>Second List Intrusions</u>				
Group(BAI vs FAI)	0.00	1	0.00	0.00
S(Group)	40.20	38	1.06	
Pair type	1.25	1	1.25	1.62
Group X Pair type	67.75	1	67.75	87.22*
S(Group) X Pair type	29.30	38	.7711	

*p<.01